PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 5/10, 15/12, G01N 33/52

A1

(11) International Publication Number: WO 93/24613

(43) International Publication Date: 9 December 1993 (09.12.93)

(21) International Application Number: PCT/US93/04707

(22) International Filing Date: 14 May 1993 (14.05.93)

(30) Priority data:
887,712 22 May 1992 (22.05.92) US

(71) Applicants: THE UNITED STATES GOVERNMENT as represented by THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Bldg. 31, Room 5A50, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20878 (US). GENETIC THERAPY, INC. [US/US]; 19 Firstfield Road, Gaithersburg, MD 20878 (US).

(72) Inventors: McDONAGH, Kevin, T. ; 10007 Greenock Road, Silver Spring, MD 20901 (US). NIENHUIS, Arthur ; 6320 Rockhurst Road, Bethesda, MD 20817 (US). TOLSTOSHEV, Paul ; 10029 Colebrook Avenue, Potomac, MD 20854 (US).

(74) Agents: OLSTEIN, Elliot, M. et al.; Carella, Byrne, Bain, Gilfillan, Cecchi & Stewart , 6 Becker Farm Road, Roseland, NJ 07068 (US).

(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: IMPROVED EXPRESSION OF HUMAN MULTIDRUG RESISTANCE GENES AND IMPROVED SELECTION OF CELLS TRANSDUCED WITH SUCH GENES

(57) Abstract

A DNA sequence for a human mdrl gene, which encodes p-glycoprotein, wherein at least one base in a splice region of the DNA encoding p-glycoprotein is changed. Such a mutation prevents truncation of the p-glycoprotein upon expression thereof. There is also provided a method of identifying cells which express the human mdrl gene in a cell population that has been transduced with an expression vehicle including a human mdrl gene. The method comprises contacting the cell population with a staining material, such as rhodamine 123, and identifying cells which express the human mdrl gene based on differentiation in color among the cells of the cell population. This method has allowed identification of retroviral producer clones facilitate mdr gene transfer into primary cells. Repopulating hematopoietic stem cells have been genetically engineered with the human mdrl gene.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SK	Slovak Republic
CI	Côte d'Ivoire	KZ	Kazakhstan	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
CS	Czechoslovakia	LK	Sri Lanka	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	UA	Ukraine
DK	Denmark	MG	Madagascar	US	United States of America
ES	Spain	ML	Mali	VN	Viet Nam
FI	Finland	MN	Mongolia		

IMPROVED EXPRESSION OF HUMAN MULTIDRUG RESISTANCE GENES AND IMPROVED SELECTION OF CELLS TRANSDUCED WITH SUCH GENES

This invention relates to the transduction of cells with human multidrug resistance (mdr) genes and the expression thereof, as well as the selection or identification of cells transduced with such genes. More particularly, this invention relates to the enhanced expression of mdr genes by transduced cells, and an improved method of selecting or identifying cells transduced with mdr genes.

Genes for multidrug resistance such as, for example, mdr1 and mutated forms thereof, encode proteins of a class known as p-glycoprotein, that confer resistance to a range of cytotoxic drugs used in the treatment of human malignancies. The introduction and expression of the mdr1 gene in human cells may serve several useful purposes in the treatment of diseases. For example, the mdr1 gene can function as a dominant selectable marker that will allow for positive selection of cells transduced by an expression vehicle, such as, for example, a retroviral vector, ex vivo or in vivo. Used in this manner, expression of the mdr1 gene in normal human cells may be useful in increasing the proportion of a target cell population that has been transformed by an expression vehicle, such as a retroviral particle. Also, expression of the mdr1 gene in normal human

cells may permit m re intensive use of chemotherapeutic drugs in the treatment of cancer. For example, expression of the mdr1 gene in human bone marrow cells could reduce the severity and duration of cytopenias following chemotherapy, which may facilitate dose intensification while reducing morbidity and mortality.

In accordance with an aspect of the present invention, there is provided human mdr1 DNA or RNA which encodes p-glycoprotein r an analogue or fragment of such DNA or RNA which encodes an iso-form of p-glycoprotein which confers multidrug resistance. The DNA or RNA includes at least one splice site, and at least one base in the splice site has been changed to a different base which thereby inactivates the splice site. Preferably, the DNA or RNA is a DNA or RNA sequence for human mdr1 which encodes p-glycoprotein wherein at least one base in a splice site of said DNA encoding p-glycoprotein is changed.

The natural human mdr1 gene contains at least one functional splice donor site and splice acceptor site. This invention provides a modification of the mdr1 gene whereby at least one base in a splice site has been changed to a different base which thereby inactivates the splice site.

The term "human mdr1 gene" as used herein means the wild type human mdr1 gene, which encodes p-glycoprotein, as well as mutated human mdr1 genes which encode what are known as iso-forms of p-glycoprotein. Such mutated genes include a mutation in which the human mdr1 DNA encodes a p-glycoprotein in which th first 23 amino acids at the N-terminal are deleted (Currier, et al., J. Biol. Chem., Vol. 264, pgs. 14376-14381 (1989)); insertional mutation(s) in the human mdr1 DNA in which amino acids have been inserted into ATP binding or utilization sites (Currier, et al; 1989); a complementary human mdr1 DNA sequence (cDNA sequence) in which codon 185, which encodes Gly 185, has been changed from GGT to GTT to encode Val 185. (PCT Application No. WO87/05943; Choi, et al., Cell, Vol. 53, pgs. 519-529 (1988);

Safa, et al., Pr c.Nat.Acad.Sci., Vol. 87, pgs. 7225-7229 (1990)); and a chimeric gene encoding a protein having adenosine deaminase added to the N-terminal of p-glycoprotein (Germann, t al., J. Biol. Ch m., Vol. 264, pgs. 7418-7424 (1989)).

Although the scope of this aspect of the present invention is not to be limited to any theoretical reasoning, Applicants have found that aberrant splicing of the mdr1 gene may occur within the coding sequence for p-glycoprotein. Such splicing, which uses cryptic splice donor and splice acceptor sites, results in the production of expression vehicles (such as, for example, recombinant retroviruses) that contain a truncated mdr1 gene, which encodes a truncated and non-functional p-glycoprotein.

Applicants have found a cryptic splice donor site having a length of 9 bases from nucleotide 760 to nucleotide 768 of the human mdr1 gene, and which has the following sequence:

CAGGTATGC

A cryptic splice acceptor site was also found, which has a length of 16 bases from nucleotide 2729 to nucleotide 2744 of the human mdr1 gene; and which has the following sequence:

ACATTTTTCCTTCAGG

Applicants have found that, by changing at least one base of the cryptic splice donor and/or acceptor sites, one may suppress aberrant splicing of the mdr1 gene and maximize the transfer of DNA encoding functional p-glycoprotein into desired cells.

In one embodiment, the at least one base which is changed is in a splice donor site. In another embodiment, the at least one base which is changed is in a splice acceptor site. In yet another embodiment, at least one base is changed in a splice donor site and at least one base is changed in a splice acceptor site.

In a preferred embodiment, the at least one base in a splic site which is changed is in the splice donor site and is at least

one of guanine (G) or thymine (T) contained in the splice donor site wherein G and T are adjacent.

In another preferred embodiment, the at least one base in a splice site which is changed is in the splice acceptor site and is at least one of adenine (A) or guanine (G) contained in the splice acceptor site wherein A and G are adjacent.

In a more preferred embodiment, at least one base is changed in the splice donor site wherein the at least one base is at least one of G or T contained in the splice donor site wherein G and T are adjacent; and at least one base is changed in the splice acceptor site wherein the at least one base is at least one of A or G contained in the splice acceptor site wherein A and G are adjacent.

In another embodiment, the at least one base in a splice site is changed such that a codon encoding an amino acid is changed to a different codon encoding the same amino acid. Preferably, the at least one base is the third, or "wobble" base in the codon.

In one embodiment, codon 139, located in the splice donor site, and which encodes Arg, is changed from AGG to AGA, which also encodes Arg. In another embodiment, codon 733, located in the splice acceptor site, and which encodes Gln, is changed from CAG to CAA, which also encodes Gln. It is also contemplated that in yet another embodiment, both codon 139 of the mdr1 gene is changed from AGG to AGA, and codon 733 of the mdr1 gene is changed from CAG to CAA.

In one embodiment, the DNA sequence is further mutated such that at least a portion of the 5' untranslated region of the DNA has been removed.

In another embodiment, the DNA sequence may be further mutated such that at least a portion of the 3' untranslated region of the DNA has been removed. In yet another embodiment, the DNA sequence is further mutated such that at least a portion

of the 5' untranslated region of the DNA and at least a portion of the 3' untranslated region of the DNA have been removed. Although Applicants do n t intend to be limited thereby, the DNA of the human mdr1 gene is quite larg and contains extended 5' and 3' untranslated regions. Such extended regions may result in the generation of a reduced titer of viral vector particles, such as retroviral vector particles, which may be engineered with such DNA. Thus, removal of at least a portion(s) of the 5' and/or 3' untranslated regions may enable one to insert such DNA into a viral vector, such as a retroviral vector, whereby one generates an increased titer of retroviral particles. Within the scope of the present invention, up to the entire 5' untranslated region and/or up to the entire 3' untranslated region may be removed.

In accordance with another aspect of the present invention, there is provided DNA (or RNA) encoding a protein which provides for multidrug resistance, wherein the DNA includes the following sequence:

CAGGTATGC

At least one base is changed to a different base. Preferably at least one base of the GT doublet is changed.

In accordance with yet another aspect of the present invention, there is provided DNA (or RNA) encoding a protein which provides for multidrug resistance, wherein the DNA includes the following sequence:

ACATTTTTCCTTCAGG

At least one base is changed to a different base. Preferably, at least one base of the AG doublet is changed.

In one embodiment, the DNA encoding a protein which provides for multidrug resistance includes a first sequence.

CAGGTATGC, and a second sequence:

ACATTTTTCCTTCAGG

At least one base in the first sequence, and at least one base in the second sequence is changed to a different base. Preferably, at least one base of the GT doublet of the first sequence is

changed, and at least one base of the AG doublet of the second sequence is changed.

The changed DNA or RNA sequence for a human mdr1 gene, or the changed DNA or RNA encoding a protein which provides for multidrug resistance, of the present invention may be cloned into any of a variety of expression vectors by genetic engineering techniques known to those skilled in the art. Such expression vectors include, but are not limited to, prokaryotic vectors, including bacterial vectors; eukaryotic vectors, such as, for example, yeast vectors and fungal vectors, and viral vectors, such as, but not limited to, retroviral vectors, and non-retroviral vectors such as, but not limited to, adenoviral vectors, adeno-associated viral vectors, and Herpes virus vectors. In one embodiment, the changed DNA (or RNA) is cloned into a retroviral expression vector. It is also contemplated that the changed DNA or RNA may be introduced into cells by means of non-viral systems and non-plasmid-based systems. Such systems include, but are not limited to, the injection of the naked changed DNA or RNA into a desired cell, liposomes which encapsulate such DNA or RNA, and which deliver such DNA or RNA to a cell, and the coupling of the DNA or RNA to a protein or other agent which binds to a cellular receptor.

Retroviral vectors which may be employed include those derived from Moloney Murine Leukemia Virus, Moloney Murine Sarcoma Virus, Rous Sarcoma Virus, Harvey Sarcoma Virus, and Spleen Necrosis Virus. In one embodiment, the retroviral vector may be derived from Moloney Murine Leukemia Virus and is one of the LN series of vectors, as described in Miller, et al., Biotechniques, Vol. 7, pgs. 980-990 (1989). Such vectors have a portion of the packaging signal derived from a mouse sarcoma virus, and a mutated gag initiation codon.

In another embodiment, the retroviral vector may include at least four cloning, or restriction enzyme recognition sites, wherein at least two of the sites have an average frequency of

appearance in eukary tic genes of less than nce in 10,000 base pairs; i.e., the restricti n product has an average DNA size of at least 10,000 base pairs. Preferred cloning sites are selected from the group c nsisting of NotI, SnaBI, SalI, and XhoI. In a preferred embodiment, the retroviral vector includes each of these cloning sites.

When a retroviral vector including such cloning sites is employed, there may also be provided a shuttle cloning vector which includes at least two cloning sites which are compatible with at least two cloning sites selected from the group consisting of NotI, SnaBI, SalI, and XhoI located on the retroviral vector. The shuttle cloning vector also includes at least one desired gene which is capable of being transferred fr m the shuttle cloning vector to the retroviral vector.

The shuttle cloning vector may be constructed from a basic "backbone" vector or fragment to which are ligated one or more linkers which include cloning or restriction enzyme recognition sites. Included in the cloning sites are the compatible, or complementary cloning sites hereinabove described. Genes and/or promoters having ends corresponding to the restriction sites of the shuttle vector may be ligated into the shuttle vector through techniques known in the art.

The shuttle cloning vector can be employed to amplify DNA sequences in prokaryotic systems. The shuttle cloning vector may be prepared from plasmids generally used in prokaryotic systems and in particular in bacteria. Thus, for example, the shuttle cloning vector may be derived from plasmids such as pBR322; pUC 18; etc.

Examples of retroviral vectors having at least two restriction enzyme recognition sites having an average frequency of appearance in eukaryotic genes of less than once in 10,000 base pairs, and of the shuttle cloning vectors employed in transfering genes to such vectors are further described in PCT Application No. WO91/10728.

The vectors also include one more m re pr moters. Suitable promoters which may be employed include, but are not limited to, the retroviral LTR; th SV40 promoter; and the human cytomegalovirus (CMV) promot r described in Miller, et al., Biotechniques, Vol. 7, No. 9, pgs. 980-990 (1989), or any other promoter (e.g., cellular promoter such as eukaryotic cellular promoters including, but not limited to, the histone, pol III, and ß-actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, TK promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art from the teachings contained herein.

The vector may also further include a heterologous or foreign gene.

Heterologous or foreign genes which may be placed into the vectors of the present invention include, but are not limited to genes which encode cytokines or cellular growth factors, such as lymphokines, which are growth factors for lymphocytes. Other examples of foreign genes include, but are not limited to, genes encoding soluble CD4, Factor VIII, Factor IX, ADA, the LDL receptor, ApoI, tumor necrosis factors (TNF's) and ApoC.

Suitable promoters which may control the foreign genes include those hereinabove described.

The vector including the DNA of the present invention may be transduced into a suitable packaging cell line. Examples of suitable packaging cell lines include the PA317 cell line and the PE501 cell line (Miller, et al., 1989), and the PAT 2.4 cell line (U.S. application Serial No. 792,281, filed November 14, 1991.) Transduction of the packaging cell line may be accomplished by standard techniques such as electroporation or $CaPO_4$ precipitation. Vector particles generated from such producer cells may then be employed to generate a producer cell line, and to transduce cells (eg., eukaryotic cells such as mammalian cells), which may be administered to a host as part of a gene

therapy procedure. Examples of cells which may be transduced include, but are not limited to, primary human cells such as primary human nucleated blood cells (such as leukocytes, and lymphocytes such as TIL cells, T-lymphocytes, and B-lymphocytes), tumor cells, endothelial cells, epithelial cells, keratinocytes, stem cells, bone marrow cells, hepatocytes, connective tissue cells, fibroblasts, mesenchymal cells, mesothelial cells, and parenchymal cells.

Upon transfection or transduction of a packaging cell line with a retroviral vector, it is essential that the producer cell line generate a sufficient titer of vector particles for gene therapy protocols. Selection for growth of such packaging or producer cells in the presence of colchicine has resulted in the isolation of producer clones with viral titers below what is currently adequate for clinical gene therapy protocols.

It is therefore another object of the present invention to provide a method of selecting producer cells which generate an acceptable titer of retroviral particles including a human mdr1 gene, and in particular for generating an adequate titer of retroviral particles suitable for inserting and expressing the mdr1 gene in primate cells, particularly in primate repopulating hematopoietic stem cells.

In accordance with an aspect of the present invention, there is provided a method of identifying cells which overexpress a gene encoding multidrug resistance (preferably such gene is the human mdr1 gene) in a cell population transduced with an expression vehicle including a gene which encodes a protein which provides for multidrug resistance (preferably such gene is the human mdr1 gene). The method comprises contacting the cell population with a staining material, and identifying the cells which overexpress the gene encoding a protein which provides for multidrug resistance (preferably the human mdr1 gene) based on differentiation of color among the cells of the cell population. Such differentiation in color may be based on differences in

color intensity between cells which ov rexpress th human mdr1 gene and cells which do not overexpress the human mdr1 gene (for example, cells which contain vectors including the human mdr1 gene, wher by such cells overexpress the human mdr1 gene, may be brighter or duller than those cells which do not contain vectors including the human mdr1 gene, whereby such cells which express the human mdr1 gene at normal levels or do not express the gene); or differences in color or shades of color between cells which contain vectors including the human mdr1 gene, and thereby overexpress the human mdr1 gene, and cells which do not contain vectors including the human mdr1 gene, and thereby express the human mdr1 gene at normal levels or do not express the human mdr1 gene. Alternatively, cells which contain vectors including the human mdr1 gene, and thereby overexpress the human mdr1 gene may change color after staining, while cells which do not contain vectors which include the human mdr1 gene do not change color after staining.

The term "overexpress" as used herein, means that the g ne encoding multidrug resistance is expressed at a level above that which is expressed by a normal human cell. In normal human cells, the mdr1 gene is expressed at low levels or is not expressed. In cells transduced with an expression vehcile including the mdr1 gene, the expression level of the mdr1 gene is considerably increased, and thus, the mdr1 gene is "overexpressed."

Applicants have found that, by identifying cells which overexpress the human mdr1 gene in a cell population transduced with infectious viral particles which include the human mdr1 gene, in accordance with the method of the present invention, Applicants have been able to isolate clones which produce higher viral titers than clones selected with drugs such as colchicine, for example.

In one embodiment, the staining material is a fluorescent dye. Fluorescent dyes which may be employed include, but are not

limited to, rhodamine 123, 3,3'-diethylc:acarbocyanine iodide ($DiOC_2$), and 3,3'-di thyloxadicarb cyanine iodide (DODC iodide).

In on emb diment, the flu resc nt dye is rhodamin 123. Rhodamine 123 ordinarily concentrates within the mitoch ndria of cells, but can be effluxed from the cell by p-glycoprotein. Cells stained with rhodamine 123 that do not express p-glycoprotein or express p-glycoprotein at normal or lower than normal levels appear "bright" by fluorescence microscopy or by FACS analysis, whereas cells that overexpress p-glycoprotein appear "dull," even if small quantities of p-glycoprotein ar expressed. (Chaudhary, et al., Cell, Vol. 66, pgs. 85-94 (July 12, 1991)). Thus, producer cells or other cells which have incorporated and express a functional human mdr1 gene may be distinguished from non-transformed cells on the basis of this "dull" staining phenotype.

Such a method may also be employed to determine viral titer of human mdr1 retrovirus producer cell lines. In one embodiment, a defined number of target cells are exposed to serial diluti ns of a viral supernatant produced by a defined number of producer cells. After infection (eg., at about 48 hrs. after infection), the cells are stained with rhodamine 123 and allowed to efflux the dye. The proportion of "dull" cells can be visually estimated by fluorescence microscopy, or accurately quantitated by scoring a defined number of individual cells by FACS analysis. Because the proportion of cells overexpressing p-glycoprotein is known, it is possible to calculate accurately the minimum number of infectious viral particles present in a known volume of viral supernatant.

Although it has been demonstrated that retroviral vectors may be used to transfer and express the mdr1 gene in cultured cells (Proc.Nat.Acad.Sci., Vol. 85, pgs. 1595-1596 (1988)); in murine repopulating hematopoietic stem cells (Sorrentino, et al., Science, (in press)), and in murine hematopoietic progenitors in vitro (McLachlin, et al., J.Nat.Canc.Inst., Vol. 82, pgs.

1260-1263 (1990)), there has been no demonstrati n of transfer and expression of the mdr1 gene in primate cells, and in particular in primate stem cells capable f rec nstituting transplant recipients.

Thus, in accordance with an aspect of the present invention, there is provided a primate cell which is genetically engineered with DNA (RNA) which encodes a protein which provides multidrug resistance. The DNA (RNA) may be any DNA (RNA) which encodes a protein which provides multidrug reistance or the changed DNA or RNA sequences which are hereinabove described.

Applicants have discovered that by selecting for cells transduced with a gene encoding multidrug resistance in accordance with the selection method hereinabove described, one is able to obtain producer cells which generate a sufficient titer of infectious viral particles which provide for the transfer and expression of multidrug resistance genes in primate cells, in particular in primate repopulating hematopoietic stem cells.

Primate cells which may be genetically engineered with such DNA (RNA) include, but are not limited to, primate bone marrow cells, preferably enriched primate bone marrow cells, and hematopoietic progenitor (CD34+) cells or hematopoietic stem cells (in particular, primate repopulating hematopoietic stem cells), and human primary cells. The hematopoietic stem cells may also be $CD33^-$ and $HLA\text{-}DR^{low}$. (Chaudhary, et al., 1991) The primate cells may be obtained from humans or other primates, such as monkeys, for example.

This aspect of the present invention is particularly applicable to the use of primate repopulating hematopoietic stem cells which are genetically engineered with DNA (RNA) which encodes a protein which provides multidrug resistance. Such stem cells give rise to bone marrow cells and may be administered, for example, to a patient who has undergone chemotherapy and/or radiation treatment in order to regenerate healthy bone marrow

cells in the patient. Thus Applicants have devised a system for identifying pr ducer cells which generate a sufficient titer of viral particles including a gene encoding multidrug resistance, whereby such viral particles may be used to infect primate repopulating hematopoietic stem cells, which may then be administered to a patient in order to regenerate bone marrow in said patient, wherein such regenerated bone marrow is resistant to the toxic effects of chemotherapy.

In one embodiment, bone marrow cells are harvested by n edle aspiration from a primate, and purified by positive selection for cells expressing the CD34 antigen. CD34 selected cells are enriched 50-100 fold in clonogenic hematopoietic progenitors and contain all of the repopulating stem cells. Alternatively, the bone marrow cells may be purified in accordance with the procedures described in U.S. Patent No. 5,061,620.

The CD34+ cells are incubated in vitro in the presence of stimulatory hematopoietic growth factors and in the presence of retroviral supernatant containing infectious viral particles which include a multidrug resistance gene. The viral particles are generated from a producer cell line selected in accordance with the selection procedure hereinabove described. After incubation, the cells are transplanted into recipients suitably treated to receive a bone marrow transplant. After transplantation, bone marrow cells, peripheral blood leukocytes, purified granulocytes, and T-lymphocytes are assayed for presence and expression of the introduced mdr1 gene.

The invention will be described with respect to the following examples; however, the scope of the present invention is not intended to be limited thereby.

EXAMPLE 1

A. Construction of plasmid pG1.

Plasmid pG1 was constructed from pLNSX (Palmer et al., Blood, 73:438-445; 1989). The construction strategy for plasmid pG1 is shown in Figure 1. The 1.6 kb EcoRI fragment, containing

the 5' Moloney Sarcoma Virus (MoMuSV) LTR, and the 3.0 kb EcoRI/ClaI fragment, c ntaining the 3' LTR, the bacterial origin of replicati n and the ampicillin resistance gene, were isolated separately. A linker c ntaining seven unique cl ning sites was then used to close the EcoRI/ClaI fragment on itself, thus generating the plasmid pGO. The plasmid pGO was used to generate the vector plasmid pG1 by the insertion of the 1.6 kb EcoRI fragment containing the 5' LTR into the unique EcoRI site of pGO. Thus, pG1 (Figure 3) consists of a retroviral vector backbone composed of a 5' portion dervied from MoMuSV, a short portion of gag in which the authentic ATG start codon has been mutated to TAG (Bender et al. 1987), a 54 base pair multiple cloning site (MCS) containing from 5' to 3' the sites EcoRI, NotI, SnaBI, SalI, BamHI, XhoI, HindIII, ApaI, and ClaI, and a 3' portion of MoMuLV from base pairs 7764 to 7813 numbered as described in (Van Beveren et al., Cold Spring Harbor, Vol. 2, pg. 567, 1985). (Figure 2). The MCS was designed to generate a maximum number of unique insertion sites, based on a screen of non-cutting restriction enzymes of the pG1 plasmid, the neo^R gene, the ß-galactosidase gene, the $hygromycin^R$ gene, and the SV40 promoter.

B. Construction of pG1MD1.

pMDR2000 (Ueda, et al., PNAS, Vol. 84, pgs. 3004-3008 (May 1987)), which contains an mdr1 cDNA sequence (Figure 4) described in PCT application number WO87/05943, wherein the first 282bp of the untranslated 5' region and the last 23bp of the untranslated 3' region of the cDNA sequence have been removed, was cut with SacI and EcoRI, and was inserted into a multiple cloning sit of SacI and EcoRI digested pGEM2 (Figure 5.) (Promega, Madison, WI) to form pGEM2MDR. pGEM2MDR was then cut with SmaI at the 5' end of the multiple cloning site and a SacII linker having a length of 8bp was inserted. At the 3' end of the plasmid, the plasmid was cut with EcoRI and a 12bp XhoI linker was inserted. The resulting plasmid was pMDR2000XS. (Figure 6.)

Plasmid pMDR2000XS was then cut at the 5' end of the cDNA sequence for mdr1 with HhaI, and at the 3' end f the cDNA sequence for mdr1 with HaeII, and a 3,845bp fragment was removed. Such cutting removed an additional 131 bp from the 5' end, and an additional 384 bp from the 3' end of the cDNA. A linker was then added to the 3' sequence to reconstitute the amino acids that were cut off by cutting with HaeII and to add the TGA stop (or termination) codon back into the sequence. The linker has four restriction enzyme sites, HpaI, XhoI, BglII, and ClaI, and has the following sequence:

```
5'-A G T G A A C T C T G G T T A A C T C C
                    10                  20
A C T C G A G C A C A G A T C T G G A
                  30
C A T C G A T A C T C
40                  50
```

Bases 12-17 are recognized by HpaI, bases 22-27 are recognized by XhoI, bases 31-36 are recognized by BglII, and bases 41-46 are recognized by ClaI.

pG1 was then cut with SnaBI and ClaI, and the mdr1 fragment containing the added linker was cut with ClaI and inserted into the SnaBI and ClaI digested pG1 to form pG1MD1 (Figure 7.)

C. Changing codons 139, 185, and 733 of the mdr1 gene.

Changing the mdr1 gene in pG1MD1 such that codon 139 is changed from AGG to AGA, codon 185 is changed from GTT to GGT (whereby Val is changed to Gly in the resulting expressed p-glycoprotein), and codon 733 is changed from CAG to CAA may be accomplished by either of the following strategies.

1. Introduction of point mutations into G1MD1 by site-directed mutagenesis.

pG1MD1 is digested with NotI and XhoI, and a NotI-XhoI fragment is cloned into the phagemid pBluescriptSK+(Stratagene). The resulting subclone is transformed into E. coli strain CJ236 (Biorad). A 5ml culture is grown at 37°C overnight in LB broth

with 200µg/ml ampicillin and 30µg/ml chloramphenicol. 2.5ml of this culture is then added to 25 ml of fresh media with the same antibiotics, and grown at 37 C until the O.D.600 of the culture reaches 0.3. Helper phage R408 (Stratagene) is then added at a multiplicity of infection of 20 to the culture. The culture is then incubated by vigorous shaking for 8 hours. The culture is then harvested by centrifugation at 17,000xg for 15 minutes at 4°C. The supernatant fluid is harvested and spun again at 17,000xg for 15 minutes at 4°C. The supernatant is then harvested and measured for volume.

A 1/4 volume of a solution containing 3.5 M ammonium acetate and 20% polyethylene glycol is then added to the supernatant to precipitate the phage. The solution and supernatant mixture is incubated at room temperature for 15 minutes. The precipitate is then centrifuged at 11,000xg for 15 minutes at 4°C. The resulting pellet is suspended in 400µlTE (10mM Tris, 1mM EDTA) and extracted with phenol/chloroform until the interphase is clear. The DNA (phage/template) is then ethanol precipitated by standard procedures.

2ng of the following antisense oligonucleotides:

(a)5'-T C T G T A C T G G T C T A T A C G G A T A A T
10 20
A A T G T C - 3'
30 , wherein the TCT sequence (nucleotides 11-13) will provide for a change of codon 139 from AGG to AGA;

(b)5'-C T A A T T A C T T C C A T A A C C A C T G T T
10 20
T T A A C C - 3'
30
, wherein the CCA sequence (nucleotides 11-13) will provide for a change of codon 185 from GTT to GGT (whereby Val 185 is changed to Gly 185); and

(c)5' - G T A A A A A G G A A G T T C C A A A G T G T A
10 20

A A C C G T - 3'

30 , wherein the GTT sequence (nucleotides 12-14) will provide for a change of codon 733 from CAG to CAA, and 200ng of the phage/template DNA are then mixed in 10µl of TE. The mixture is heated to 70°C, and then allowed to cool slowly to room temperature in order to anneal the oligonucleotides to the template. Deoxynucleotide triphosphates are then added to the mixture, followed by T4 DNA polymerase, T4 ligase, 10X synthesis buffer, and water to make the volume 20µl. The mixture is then placed on ice for 5 minutes, then at 25°C for 5 minutes, and then at 37°C for 90 minutes.

At the end of the 90 minute incubation, 90µl of TE is added and 5µl of the resulting mixture is used to transform E. coli strain DH5α.

Mutagenized clones are then identified by colony hybridization with the oligonucleotides (a), (b), and (c) hereinabove described. The identified and changed mdr1 gene is then cut out of pBluescript with NotI and XhoI and cloned into pG1 cut with NotI and XhoI to form pG1MD3.

2. Introduction of Point Mutations into G1MD1 by PCR Based Strategy

The cryptic splice donor site and the Gly to Val point mutation in codon 185 are contained within a NotI to ApaI restriction fragment of pG1MD1. This NotI to ApaI fragment can be excised, purified, and subcloned into a pUC based cloning vector called pUC007 (Sorrentino, et al., Nucleic Acids Research, Vol. 18, No. 9, pgs. 2721-2731 (1990)) that contains a synthetic polylinker multiple cloning site. The vector is prepared by partial digestion of pUC007 with FspI, and complete digestion with ApaI. The insert is prepared by digesting with NotI, blunting the NotI digested end with Klenow, and then digesting with ApaI. Following ligation, the abutment of the blunt FspI site with the blunted NotI site will recreate a NotI site that can be used to excise a NotI to ApaI fragment.

Following subcloning f the NotI to ApaI fragment into pUC007 (to form pMD1NA), the cryptic splice don r site and the codon 185 point mutati n will be c ntained within a uniqu BstXI to MscI fragment. A strategy of overlapping or "rec mbinant" PCR to introduce point mutations into this BstXI to MscI fragment is employed. (Higuchi, PCR Protocols: A Guide to Methods and Applications, Innis, et al., eds, San Diego, Academic Press, pgs. 177-183 (1990)). Three sets of PCR primers are made. One set of complementary primers (SD-A and SD-B), which have the following sequences:

SD-A:GGAAGACATGACCAGATATGCCTATTATTACAG

SD-B:CTGTAATAATAGGCATATCTGGTCATGTCTTCC

are centered around the cryptic splice donor site, and differ from the sequence of G1MD1 at only one base, located in codon 139 of the mdr1 gene, wherein AGG is changed to AGA. This difference inactivates the splice donor site, although the amino acid is not changed. The second set of primers (C185-A and C185-B), which have the following sequences:

C185-A:CTCTAAGATTAATGAAGGTATTGGTGACAAAATTG

C185-B:CAATTTTGTCACCAATACCTTCATTAATCTTAGAG

are centered at the codon 185 mutation of the mdr1 gene and will differ from the sequence of G1MD1 only at the point mutation wherein codon 185 is changed from GTT to GGT. This difference corrects the point mutation in the cDNA sequence (Figure 7) and encodes Gly in the final construction. The third set of primers flank the BstXI and MscI restriction sites. These primers (5'Bst and 3'Msc), which have the following sequences:

5'Bst:ATCGCGGATCCATGGTGGTGGGAACTTTGGC

3'Msc:CATCCGGAATTCAGCTGACAGTCCAAGAACAGGACTGATG

overlap the BstXI and MscI sites, respectively, and also incorporate another flanking restriction site that will allow the PCR fragment to be conveniently subcloned. Using the techniqu of "recombinant" PCR, three separate PCR reactions will be initiated, using MD1NA as the template DNA. In these reactions,

a "proofr ading" therm stable DNA polymerase, such as Vent polymeras marketed by New England BioLabs, is us d to avoid misincorporation of nucleotides, as can occur with high frequency when Taq polymerase is used. In the first PCR reaction, the two primers are 5'Bst and SD-B, and amplify the sequence from the BstXI site to the cryptic splice donor site. In the second PCR reaction, the two primers are SD-A and C185-B, and amplify the sequence from the splice donor to the codon 185 mutation. In the third PCR reaction, the primers are C185-A and 3'Msc, and amplify the sequence from the codon 185 mutation to the MscI site. Following ten cycles of amplification, the product of these three reactions is purified away from unused primers, nucleotides, and buffer. The purified PCR products are mixed together, and then amplified for an additional 10 to 20 cycles with a "proofreading" polymerase such as Vent polymerase, using the 5'Bst and 3'Msc primers. This final PCR reaction results in an amplified PCR product that extends from the BstXI site to the MscI site, and has introduced point mutations into the cryptic splice donor sit and the mutant codon 185. Using the flanking restriction sites, the PCR product is digested with BamHI and EcoRI and is subclon d into a pUC based vector prepared by digestion with BamHI and EcoRI using inactivation of the lacZ gene to identify vectors that have incorporated the insert. This construct is referred to as pMD1BM and is used to transform bacteria. Bacteria that contain the plasmid carrying the desired point mutations will be identified by hybridization to allele specific oligonucleotides. MD1BM plasmid DNA isolated from these subclones is then sequenced, using standard M13 sequence primers, to confirm that the desired mutations have been introduced, and that no additional mutations have been created.

After sequencing, a BstXI to MscI fragment is excised from MD1BM and subcloned into a vector prepared by digestion of MD1NA with BstXI and MscI. This plasmid is referred to as MD1NA-C. A NotI to ApaI fragment is excised from MD1NA-C and inserted into a

vector prepard by digestion of G1MD1 with NotI and ApaI. The resulting plasmid is referred to as pG1MD1A, and contains a cDNA for a human mdr1 gene that has a conservative point mutation inactivating a cryptic splice donor site, as well as another point mutation that causes a reversion in the amino acid sequence at codon 185 to the wild type sequence.

The cryptic splice acceptor site can be corrected using a similar PCR based strategy. A HindIII to XhoI fragment containing the cryptic splice acceptor can be subcloned directly into the HindIII and XhoI sites of pUC007 (This plasmid is designated pMD1HX). The cryptic splice acceptor is contained within an XmnI to KpnI fragment. Recombinant PCR is then us d to introduce a point mutation into the cryptic splice acceptor site. One primer set is required. The first primer (5'Xmn) having the following sequence:

5'Xmn:

ATCGCGGATCCGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGG

overlaps the XmnI site and extends past the cryptic splice acceptor. This primer will be identical to the sequence in G1MD1 except at one base in codon 733 wherein codon 733 is changed from CAG to CAA, which inactivates the splice acceptor site. The amino acid encoded by codon 733, however, is unchanged. This primer also incorporates a flanking BamHI site to facilitate subcloning of the PCR product. The second primer (3'Kpn), having the following sequence:

3'Kpn:

CTCAAAGAGTTTCTGTATGGTACC

overlaps the KpnI site. These primers are used to amplify a fragment of DNA from pMD1HX with a proofreading polymerase such as Vent polymerase. This PCR product is digested with BamHI and KpnI and subcloned into the BamHI and KpnI sites of a cloning vector such as pUC19 (This plasmid is designated pMD1XK). This plasmid is used to transform bacteria. Subclones containing the "corrected" sequence are identified by allele specific

oligonucleotide hybridizati n and sequencing as described above. Subsequently, an XmnI to KpnI fragment from pMD1XK is subcloned into the XmnI and KpnI sites of pMD1HX (The resulting plasmid is designated pMD1HX-C). Finally, a HindIII to XhoI fragment from MD1HX-C is subcloned into the HindIII and XhoI sites of pG1MD1 to form pG1MD1B, which contains a cDNA for the mdr1 gene that has a conservative point mutation inactivating a cryptic splice acceptor site.

In order to construct a plasmid that contains a cDNA for an mdr1 gene with mutations introduced into the cryptic splice donor site, codon 185, and the cryptic splice acceptor site, a HindIII to XhoI fragment is excised from pG1MD1B. This fragment is then subcloned into a vector prepared by digesting pG1MD1A with HindIII and XhoI. The resulting plasmid is referred to as pG1MD1AB.

D. Generation of producer cell lines from pG1MD3 and pG1MD1AB.

The vectors pG1MD3 and pG1MD1AB, described in Example 1, which contain the changes in codon 139 (in the splice donor site), in codon 733 (in the splice acceptor site), and in codon 185 (wherein Val 185 is changed to Gly 185), may be placed into a packaging cell line to generate vector particles, and to generate producer cell lines.

Vector producer cell lines are prepared using established protocols. The packaging cell line PE501 (Miller and Rosman, Biotechniques 7:980-990 (1989)) or cell line GP+E86 is plated at a density of 5×10^5 cells per 100 mm plate and the following day purified vector DNA is introduced using standard $CaPO_4$ precipitation (Wigler et al., Cell 14725-731 (1978)). For each plate of cells to be transfected, 20-40 µg of vector DNA is prepared with a co-precipitate consisting of 0.25M $CaCl_2$/1 mM Hepes (pH 7.2) and 140 mM NaCl, 0.75 mM Na_2HPO_4, 25 mM Hepes (pH7.2). The DNA/precipitate is allowed to sit at room temperature for 30 min and then added (1 ml/plate) to the cells

in tissue culture medium (DMEM + 10% f tal Bovin serum) for an overnight incubation. The medium is changed to fresh DMEM + serum the f llowing morning. The transfected cells are allowed to grow to near confluence for the next 48 hours, at which point virus supernatant is collected to infect a separate population of PA317 or GP+AM12 vector packaging cell lines at a density of 1 x 10^5 cells per 100 mm plate seeded 24 hours prior to infection. The standard infection conditions include undiluted virus supernatant, filtered through a 0.2 uM membrane, to which 8 ug/ml polybrene is added. After overnight incubation for 16 hours, the medium is changed to DMEM and 10% Fetal Bovine Serum, and grown until such cells are selected with rhodamine 123.
The polyclonal population of GP+Am12 or PA 317 cells are incubated with rhodamine 123 at a concentration of 1 microgram per ml at 37°C for 30 minutes. Following staining, the cells are allowed to efflux the dye for two hours in medium that did not contain rhodamine. The cells are then trypsinized, and 5 x 10^6 cells are sorted on a Coulter EPICS Elite FACS machine, with excitation at 480nm and fluorescence measured at 525 nm. The bright cells, which did not express p-glycoprotein, are discarded. Cells that are "dull" following rhodamine staining express p-glycoprotein and are collected for further processing.

Example 2

Selection of producer cells generating high titers of viral particles including an mdr1 gene.

PA317 (amphotropic) packaging cells were transfected by the $CaPO_4$ method with pG1MD1. 48 hours later, supernatant from the transfected PA317 cells (Miller, et al., 1989) was filtered and added to GP + E86 (ecotropic) packaging cells in the presence of 6 micrograms per ml of polybrene. 24 hours following infection with this transient supernatant, drug selection with colchicine at a concentration of 60 ng/ml was applied. Afte two weeks of drug selection, individual drug resistant GP + E86 clones were isolated and characterized for production of recombinant

retrovirus using an RNA sl t blot technique (Bodine, et al., Proc. Nat. Acad. Sci., Vol. 87, pgs. 3738-3742 (May 1990)). The ecotropic MDR producer clone with the highest apparent titer, referred to as E4, was used in the subsequent experiments.

Filtered viral supernatant from the ecotropic E4 producer cell line was used to infect the amphotropic packaging cell line GP + Aml2 (Markowitz, et al., Virology, Vol. 167, pgs. 400-406 (1988), Markowitz, et al., J. Virol., Vol. 62, pgs. 1120-1124 (1988)) In this particular experiment, the GP + Aml2 line was exposed to a total of 6 viral supernatants over the course of one week. Following infection, GP + Aml2 cells that had been transduced with the MDR virus and expressed p-glycoprotein were selected by FACS using rhodamine 123 staining to identify transduced cells. The polyclonal population of GP + Aml2 cells was incubated with rhodamine 123 at a concentration of 1 microgram/ml at 37 degrees C for 30 minutes. Following staining, the cells were allowed to efflux the dye for two hours in medium that did not contain rhodamine. The cells were then trypsinized, and 5 x 10^6 cells were sorted on a Coulter EPICS Elite FACS machine, with excitation at 480nm and fluorescence measured at 525nm. The bright cells, that did not express p-glycoprotein, were discarded. Cells that were "dull" following rhodamine staining express p-glycoprotein and were collected for further processing. In this particular experiment, the "dull" population was further subdivided into three groups ("A", "B", and "C") based on the degree of "dullness." These dull cells were plated at limiting dilution. A total of 1500 wells were plated at two concentrations of one cell per two wells and one cell per five wells. Following two weeks in culture, individual subclones were expanded and characterized for estimated viral titer using the RNA slot blot technique on filtered viral supernatant. Ten clones were identified with the highest production of packaged and secreted viral RNA. These clones were subsequently characterized for viral titer using the rhodamine titering

protocol. In this protocol, 10 ml of media is conditioned by 5 x 10^6 pr ducer cells for 24 hours. The supernatant is filtered, and dilutions of the filtered supernatant are used to infect a target cell population, in this case 3T3 cells. To infect 3T3 cells a total of 10 ml of media containing viral supernatant and 6 micrograms per ml of polybrene (or 5 micrograms per ml of protamine sulfate) is added to 10^6 target cells on a 10 cm plate for a period of 48 to 72 hours. The plate of infected 3T3 cells is then stained with rhodamine 123 at 1 microgram/ml at 37 degrees for 30 minutes, and then destained for two hours in media without rhodamine. The plate is then trypsinized and analyzed by FACS to quantitate the proportion of cells that express the "dull" phenotype. Alternatively, the plate of cells can be directly examined by fluorescence microscopy. This information is then used to calculate viral titer. For example, if 0.1 ml of viral supernatant is able to confer the "dull" phenotype to 10 percent of 10^6 cells, then 0.1 ml of viral supernatant contains a minimum of 1x 10^5 viral particles, and 1 ml of viral supernatant contains a minimum of 1 x 10^6 viral particles. As shown in Figure 8A, as the volume of viral supernatant from producer clone G1MD1 A1.2 which contacts the 3T3 cells increases from 0.025 ml to 0.1 ml, and then to 1.0 ml, the height of the left peak, which corresponds to the number of cells having the "dull" phenotype, and which therefore have a lower fluorescence intensity, increases, while there is a decrease in the number of cells having the "bright" phenotype (such cells have a higher fluorescence intensity) as evidenced by the decrease in the size of the right peak. Using this assay, a identified GP + Am clone with a titer of 4 x 10^6 viral particles per ml was identified. This producer clone was designated G1MD1 A1.2.

Example 3

PA317 cells (Miller, et al., 1989), obtained from the ATCC and selected in HMT medium (following the protocol suggested by the ATCC) were transduced with viral supernatant from the

ecotropic E4 producer cell line obtained as hereinabove described in Example 2. The PA317 packaging cell line was serially infected with E4 viral supernatant over th course of 10 days until 100% of the cells in the population express d the rhodamine "dull" phenotype. The phenotype was evaluated by fluorescence microscopy and FACS analysis on aliquots of the population. A total of 13 exposures to viral supernatants were used. The population of infected PA317 cells were directly plated at limiting dulution. A total of 1,500 wells were plated at dilutions of 1:3 and 1:5 cells per well. Individual clones were isolated and subsequently scored for viral RNA production using the slot blot assay. The subclones with the highest production of viral RNA were then analyzed by the rhodamine titering assay hereinabove described in Example 2. As shown in Figure 8B, as the volume of viral supernatant from producer clone PA15 which contacts the 3T3 cells increases from 0.025 ml to 0.1 ml and then to 1.0 ml, the height of the left peak, which corresponds to the number of cells having the "dull" phenotype, and which therefore have a lower fluorescence intensity, increases, while there is a decrease in the number of cells having the "bright" phenotype (such cells have a higher fluorescence intensity) as evidenced by the decrease in the size of the right peak. One subclone, PA15, was identified which had a titer of 5×10^6 viral particles per ml.

Example 4

Bone marrow cells were harvested from the posterior iliac crests and femurs of 3-4 kg juvenile rhesus monkeys (Proc. Natl. Acad. Sci. USA, Vol. 87, pgs. 3738-3742 (1990)). The cells are aspirated into Dulbecco's Modified Eagle's Medium (DMEM) containing 2% fetal calf serum and 10 units of heparin/ml. Five days before the bone marrow harvest, the animals receive 5-fluorouracil (70mg/kg) as a single intravenous bolus. During the bone marrow harvest procedure, the animal receives 100ml of autologous red cells to replace the volume of blood that is

removed. Prior to the bon marrow harvest, an in-dwelling c ntral venous catheter is established by standard surgical techniques.

After harvest, the bone marrow was dilut d with Hank's Buffered Saline Solution (HBSS) and passed through wire mesh to remove clumps of cells. Mononuclear cells were isolated by density gradient centrifugation using Ficoll-Paque. Cells expressing the CD34 antigen are isolated by immunoselection (Science, Vol. 242, pgs. 919-922 (1988)) using the magnetic bead technology. $2\text{-}4X10^9$ total bone marrow cells are recovered by the harvest procedure. Approximately 20% of these cells are recovered as the mononuclear cell preparation and 2-3% of these cells are recovered by immunoselection. Thus a total of $2\text{-}4X10^7$ CD34 cells are obtained at the time of bone marrow harvest. These were incubated at an initial concentration of $5X10^5$ cells/ml for 36 hours in DMEM, 15% fetal calf serum, penicillin-streptomycin, and 2X glutamine. The following hematopoietic growth factors were added: Stem Cell Factor (100ng/ml), interleukin-3(50ng/ml) and interleukin-6 (50ng/ml). After 36 hours, the cells were recovered by centrifugation and resuspended at $5X10^5$ cells/ml in fresh medium conditioned by the retroviral producer clone G1MD1 A1.2 prepared as described in Example 2. This medium includes hematopoietic growth factors as specified above plus protamine sulfate at a concentration of 5ug/ml. Every twelve hours the cells were recovered by centrifugation and resuspended in fresh virus conditioned medium with growth factors and protamine. After 72 hours of incubation in the presence of retroviral vector particles, the cells were recovered by centrifugation, resuspended in DMEM or HBSS containing 2% fetal calf serum and 10 units of heparin per ml, and reinfused into the transplant recipient. During the in vitro culture, the cells expand approximately 10 fold so that $2\text{-}4X10^6$ cells are returned to the animal.

Initially, cells w re analyzed after in vitro incubation for evidence of gene transfer and expression. DNA was recovered and subjected to analysis by the polymeras chain r action methodology using primers (MDR Primer 7, which has the sequence 5'-GCCCACATCATCATGATC-3' and MDR Primer 8, which has the sequence 5'-GTCTCCTACTTTAGTGCT-3') specific for the human mdr1 coding sequences. Based on comparison to signal intensity obtained with a retorviral producer clone containing a single copy of the retroviral genome, we concluded that 10% of the cells had undergone transduction. The Rhodamine efflux assay hereinabove described was performed confirming that, when compared to a mock infected control sample, 10% of the CD34 selected, transduced cells expressed the mdr1 gene product.

The transplant recipients received daily intravenous fluids, broad spectrum antibiotics, antifungal agents, hyperalimentation solution and blood products as needed to maintain an optimal physiological condition. Daily blood counts and blood chemistries were obtained to monitor recovery. On day 1 following transplantation, granulocyte colony-stimulating factor was given at a dose of 5ug/kg/day by continuous intravenous infusion. When recovery to a leukocyte count of 1,000-3,000/mm^3 occurs (day 14-22), blood samples are obtained for detection of the mdr1 gene and its expression.

DNA was purified from the total population of blood leukocytes after lysis of the red cells, and isolation of nuclei from leukocytes. This DNA was analyzed by the polymerase chain reaction methodology using the P7 and P8 primers described above. In the first animal transplanted, approximately 1% of the cells contained the retroviral genome whereas in the second animal 8% of the cells contained the transferred sequences.

It is to be understood, however, that the scope of the present invention is not to be limited to the specific embodiments described above. The invention may be practiced

other than as particularly described and still be within the scope of the accompanying claims.

PATAP859

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: McDonagh, Kevin T.
Nienhuis, Arthur
Tolstoshev, Paul

(ii) TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN MULTIDRUG RESISTANCE GENES AND IMPROVED SELECTION OF CELLS TRANSDUCED WITH SUCH GENES

(iii) NUMBER OF SEQUENCES: 18

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Carella, Byrne, Bain, Gilfillan, Cecchi & Stewart
(B) STREET: 6 Becker Farm Road
(C) CITY: Roseland
(D) STATE: New Jersey
(E) COUNTRY: USA
(F) ZIP: 07068

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch diskette
(B) COMPUTER: IBM PS/2

(C) OPERATING SYSTEM: PC-DOS
(D) SOFTWARE: DW4.V2

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 07/887,712
(B) FILING DATE: 22-MAY-1992
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(2) INFORMATION FOR SEQ ID NO:1:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CAGGTATGC 9

(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: G nomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ACATTTTTCC TTCAGG 16

(2) INFORMATION FOR SEQ ID NO:3:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGTGAACTCT GGTTAACTCC ACTCGAGCAC AGATCTGGAC ATCGATACTC 50

(2) INFORMATION FOR SEQ ID NO:4:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TCTGTACTGG TCTATACGGA TAATAATGTC 30

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5

CTAATTACTT CCATAACCAC TGTTTTAACC 30

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTAAAAAGGA AGTTCCAAAG TGTAAACCGT 30

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGAAGACATG ACCAGATATG CCTATTATTA CAG 33

(2) INFORMATION FOR SEQ ID NO:8:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTGTAATAAT AGGCATATCT GGTCATGTCT TCC 33

(2) INFORMATION FOR SEQ ID NO:9:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTCTAAGATT AATGAAGGTA TTGGTGACAA AATTG 35

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: singular

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CAATTTTGTC ACCAATACCT TCATTAATCT TAGAG 35

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: singular

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATCGCGGATC CATGGTGGTG GGAACTTTGG C 31

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: singular

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CATCCGGAAT TCAGCTGACA GTCCAAGAAC AGGACTGATC 40

(2) INFORMATION FOR SEQ ID NO:13:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATCGCGGATC CGGAATTATT TCTTTTATTA CATTTTTCCT TCAAGGTTTC ACATTTGG 58

(2) INFORMATION FOR SEQ ID NO:14:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CTCAAAGAGT TTCTGTATGG TACC 24

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GCCCACATCA TCATGATC 18

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

(A) DESCRIPTION: DNA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GTCTCCTACT TTAGTGCT 18

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs, having 4 unpaired bases at the 5' end, and 2 unpaired bases at 3' end.

(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: Plasmid DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AATTCGCGGC CGCTACGTAG TCGACGGATC CCTCGAGAAG CTTGGGCCCA T
GCGCCG GCGATGCATC AGCTGCCTAG GGAGCTCTTC GAACCCGGGT AGC 53

(2) INFORMATION FOR SEQ ID NO:18:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4669 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singular
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE:
(A) DESCRIPTION: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CATACTCTAT TCAGATATTC TCCAGATTCC TAAAGATTAG AGATCATTTC TCATTCTCCT 60
AGGAGTACTC ACTTCAGGAA GCAACCAGAT AAAAGAGAGG TGCAACGGAA GCCAGAACAT 120
TCCTCCTGGA AATTCAACCT GTTTCGCAGT TTCTCGAGGA ATCAGCATTC AGTCAATCCG 180
GGCCGGGACC AGTCATCTGT GGTGAGGCTG ATTGGCTGGG CAGGAACAGC GCCGGGGCGT 240
GGGCTGAGCA CAGCGCTTCG CTCTCTTTGC CACAGGAAGC CTGAGCTCAT TCGAGTAGCG 300
GCTCTTCCAA GCTCAAAGAA GCAGATCCCG CTGTTCGTTT CCTTTAGGTC TTTCCACTAA 360
AGTCGGAGTA TCTTCTTCCA AGATTTCACG TCTTGGTGGC CGTTCCAAGG AGCGCGAGGT 420

CGGGATGGAT CTTGAAGGGG ACCGCAATGG AGGAGCAAAG AAGAAGAACT TTTTTAAACT 480

GAACAATAAA AGTGAAAAAG ATAAGAAGGA AAAGAAACCA ACTGTCAGTG TATTTTCAAT 540

GTTTCGCTAT TCAAATTGGC TTGACAAGTT GTATATGGTG GTGGGAACTT TGGCTGCCAT 600

CATCCATGGG GCTGGACTTC CTCTCATGAT GCTGGTGTTT GGAGAAATGA CAGATATCTT 660

TGCAAATGCA GGAAATTTAG AAGATCTGAT GTCAAACATC ACTAATAGAA GTGATATCAA 720

TGATACAGGG TTCTTCATGA ATCTGGAGGA AGACATGACC AGGTATCCCT ATTATTACAG 780

TGGAATTGGT GCTGGGGTGC TGGTTGCTGC TTACATTCAG GTTTCATTTT GGTGCCTGGC 840

AGCTGGAAGA CAAATACACA AAATTAGAAA ACAGTTTTTT CATGCTATAA TGCGACAGGA 900

GATAGGCTGG TTTGATGTGC ACGATGTTGG GCAGCTTAAC ACCCGACTTA CAGATGATGT 960

CTCTAAGATT AATGAACTTA TTGGTGACAA AATTGGAATG TTCTTTCAGT CAATGGCAAC 1020

ATTTTTCACT GGGTTTATAG TAGGATTTAC ACGTGGTTGG AAGCTAACCC TTGTGATTTT 1080

GGCCATCAGT CCTGTTCTTG GACTGTCAGC TGCTGTCTGG GCAAAGATAC TATCTTCATT 1140

TACTGATAAA GAACTCTTAG CGTATGCAAA AGCTGGAGCA GTAGCTGAAG AGGTCTTGGC 1200

AGCAATTAGA ACTGTGATTG CATTTGGAGG ACAAAAGAAA GAACTTGAAA GGTACAACAA 1260

AAATTTAGAA GAAGCTAAAA GAATTGGGAT AAAGAAAGCT ATTACAGCCA ATATTTCTAT 1320

AGGTGCTGCT TTCCTGCTGA TCTATGCATC TTATGCTCTG GCCTTCTGGT ATGGGACCAC 1380

CTTGGTCCTC TCAGGGGAAT ATTCTATTGG ACAAGTACTC ACTGTATTCT TTTCTGTATT 1440

AATTGGGGCT TTTAGTGTTG GACAGGCATC TCCAAGCATT GAAGCATTTG CAAATGCAAG 1500

AGGAGCAGCT TATGAAATCT TCAAGATAAT TGATAATAAG CCAAGTATTG ACAGCTATTC 1560

GAAGAGTGGG CACAAACCAG ATAATATTAA GGGAAATTTG GAATTCAGAA ATGTTCACTT 1620

CAGTTACCCA TCTCGAAAAG AAGTTAAGAT CTTGAAGGGC CTGAACCTGA AGGTGCAGAG 1680

TGGGCAGACG GTCCCCCTGG TTGGAAACAG TGGCTGTGGG AAGAGCACAA CAGTCCAGCT 1740

GATGCAGAGG CTCTATGACC CCACAGAGGG GATGGTCAGT GTTCATGGAC AGGATATTAG 1800

GACCATAAAT GTAAGGTTTC TACGGGAAAT CATTGGTGTG GTGAGTCAGG AACCTGTATT 1860

GTTTGCCACC ACGATAGCTG AAAACATTCG CTATGGCCGT GAAAATGTCA CCATGGATGA 1920

GATTGAGAAA GCTGTCAAGG AAGCCAATGC CTATGACTTT ATCATGAAAC TGCCTCATAA 1980

ATTTGACACC CTGGTTGGAG AGAGAGGGCC CCAGTTGAGT GGTGGGCAGA AGCAGAGGAT 2040

CGCCATTGCA CGTGCCCTGG TTCGCAACCC CAAGATCCTC CTGCTGGATG AGGCCACGTC 2100

AGCCTTGGAC ACAGAAAGCG AAGCAGTGGT TCAGGTGGCT CTGGATAAGG CCAGAAAAGG 2160

TCGGACCACC ATTGTGATAG CTCATCGTTT GTCTACAGTT CGTAATGCTG ACGTCATCGC 2220

TGGTTTCGAT GATGGAGTCA TTGTGGAGAA AGGAAATCAT GATGAACTCA TGAAAGAGAA 2280

AGGCATTTAC TTCAAACTTG TCACAAACA GACAGCAGGA AATGAAGTTG AATTAGAAAA 2340

TGCAGCTGAT GAATCCAAAA GTGAAATTGA TGCCTTGGAA ATGTCTTCAA ATGATTCAAG 2400

ATCACGTCTA ATAAGAAAAA GATCAACTCG TAGGAGTGTC CGTGGATCAC AACCCCAAGA 2460

CAGAAAGCTT AGTACCAAAG AGGCTCTGGA TGAAAGTATA CCTCCAGTTT CCTTTTGGAG 2520

GATTATGAAG CTAAATTTAA CTGAATGGCC TTATTTTGTT GTTGGTGTAT TTTGTGCCAT 2580

TATAAATGGA GGCCTGCAAC CAGCATTTGC AATAATATTT TCAAAGATTA TAGGGGTTTT 2640

TACAAGAATT GATGATCCTG AAACAAAACG ACAGAATACT AACTTGTTTT CACTATTGTT 2700

TCTAGCCCTT GGAATTATTT CTTTTATTAC ATTTTTCCTT CAGGGTTTCA CATTTGGCAA 2760

AGCTGGAGAG ATCCTCACCA AGCGGCTCCG ATACATGGTT TTCCGATCCA TGCTCAGACA 2820

GGATGTGAGT TGGTTTGATG ACCCTAAAAA CACCACTGGA GCATTGACTA CCAGGCTCGC 2880

CAATGATGCT GCTCAAGTTA AAGGGGCTAT AGGTTCCAGG CTTGCTGTAA TTACCCAGAA 2940

TATAGCAATT CTTGGGACAG GAATAATTAT ATCCTTCATC TATGGTTGGC AACTAACACT 3000

GTTACTCTTA GCAATTGTAC CCATCATTGC AATAGCAGGA GTTGTTGAAA TGAAAATGTT 3060

GTCTGGACAA GCACTGAAAG ATAAGAAAGA ACTAGAAGGT GCTGGGAAGA TCGCTACTGA 3120

AGCAATAGAA AACTTCCGAA CCGTTGTTTC TTTGACTCAG GAGCAGAAGT TTGAACATAT 3180

GTATGCTCAG AGTTTGCAGG TACCATACAG AAACTCTTTG AGGAAAGCAC ACATCTTTGG 3240

AATTACATTT TCCTTCACCC AGGCAATGAT GTATTTTTCC TATGCTGGAT GTTTCCGGTT 3300

TGGAGCCTAC TTGGTGGCAC ATAAACTAAT GAGCTTTGAG GATGTTCTGT TAGTATTTTC 3360

AGCTGTTGTC TTTGGTGCCA TGGCCGTGGG GCAAGTGAGT TCATTTGCTC CTGACTATGC 3420

CAAAGCCAAA ATATCAGCAG CCCACATCAT CATGATCATT GAAAAAACCC CTTTGATTGA 3480

CAGCTACAGC ACGGAAGGCC TAATGCCGAA CACATTGGAA GGAAATGTCA CATTTGGTGA 3540

AGTTGTATTC AACTATCCCA CCCGACCGGA CATCCCAGTC CTTCAGGGAC TGAGCCTGGA 3600

GGTGAAGAAG GGCCAGACGC TGGCTCTGGT GGGCAGCAGT GGCTGTGGGA AGAGCACAGT 3660

GGTCCAGCTC CTGGAGCGGT TCTACGACCC CTTGGCAGGG AAAGTGCTGC TTGATGGCAA 3720

AGAAATAAAG CGACTGAATG TTCAGTGGCT CCGAGCACAC CTGGGCATCG TGTCCCAGGA 3780

GCCCATCCTG TTTGACTGCA GCATTGCTGA GAACATTGCC TATGGACACA ACAGCCGGGT 3840

GGTGTCACAG GAAGAGATCG TGAGGGCAGC AAAGGAGGCC AACATACATG CCTTCATCGA 3900

GTCACTGCCT AATAAATATA GCACTAAAGT AGGAGACAAA GGAACTCAGC TCTCTGGTGG 3960

CCAGAAACAA CGCATTGCCA TAGCTCGTGC CCTTGTTAGA CAGCCTCATA TTTTGCTTTT 4020

GGATGAAGCC ACGTCAGCTC TGGATACAGA AAGTGAAAAG GTTGTCCAAG AAGCCCTGGA 4080

CAAACCCAGA GAAGGCCGCA CCTGCATTGT GATTGCTCAC CGCCTGTCCA CCATCCAGAA 4140

TGCAGACTTA ATAGTGGTGT TTCAGAATGG CAGAGTCAAG CAGCATGGCA CGCATCAGCA 4200

GCTGCTGGCA CAGAAAGGCA TCTATTTTTC AATGGTCAGT GTCCAGCCTG GAACAAAGCG 4260

CCAGTGAACT CTGACTGTAT GAGATGTTAA ATACTTTTTA ATATTTGTTT AGATATGACA 4320

TTTATTCAAA GTTAAAAGCA AACACTTACA GAATTATGAA GAGGTATCTG TTTAACATTT 4380

CCTCACTCAA CTTCAGAGTC TTCAGAGACT TCGTAATTAA AGGAACAGAG TGAGAGACAT 4440

CATCAAGTGG AGAGAAATCA TAGTTTAAAC TGCATTATAA ATTTTATAAC AGAATTAAAG 4500

TAGATTTTAA AAGATAAAAT GTGTAATTTT GTTTATATTT TCCCATTTGG ACTGTAACTG 4560

ACTGCCTTGC TAAAAGATTA TAGAAGTAGC AAAAAGTATT GAAATGTTTG CATAAAGTGT 4620

CTATAATAAA ACTAAACTTT CATGTGAAAA AAAAAAAAAA AAAAAAAAA 4669

WHAT IS CLAIMED IS:

1. Human mdr1 DNA or RNA which encodes p-glycoprotein or an analogue or fragment of such DNA or RNA which encodes an iso-form of p-glycoprotein which confers multidrug resistance, said DNA or RNA including at least one splice site, wherein at least one base in the splice site has been changed to a different base which thereby inactivates the splice site.

2. The DNA or RNA of Claim 1 wherein said DNA or RNA is a DNA or RNA sequence for a human mdr1 gene which encodes p-glycoprotein wherein at least one base in a splice site of said DNA encoding p-glycoprotein is changed.

3. The DNA sequence of Claim 2 wherein said at least one base is in a splice donor site.

4. The DNA sequence of Claim 2 wherein said at least one base is in a splice acceptor site.

5. The DNA sequence of Claim 2 wherein at least one base in a splice donor site is changed and at least one base in a splice acceptor site is changed.

6. The DNA sequence of Claim 2 wherein said at least one base is in a splice donor site, and is at least one of guanine or thymine contained in said splice donor site, and wherein said guanine and thymine are adjacent.

7. The DNA sequence of Claim 2 wherein said at least one base is in a splice acceptor site, and is at least one of adenine or guanine contained in said splice acceptor site, and wherein said adenine and guanine are adjacent.

8. The DNA sequence of Claim 2 wherein said at least one base in a splice site is changed such that a codon encoding an amino acid is changed to a different codon encoding the same amino acid.

9. The DNA sequence of Claim 8 wherein codon 139 is changed from AGG to AGA.

10. The DNA sequence of Claim 8 wherein codon 733 is changed from CAG to CAA.

11. The DNA sequence of Claim 8 wherein codon 139 is changed from AGG to AGA and codon 733 is changed from CAG to CAA.

12. DNA encoding a protein which provides for multidrug r sistance wherein said DNA includ s the following sequence:

CAGGTATGC

wherein at least one base is changed to a different base.

13. The DNA of Claim 12 wherein at least one base of the GT doublet is changed.

14. DNA encoding a protein which provides for multidrug resistance wherein said DNA includes the following sequence:

ACATTTTTCCTTCAGG

wherein at least one base is changed to a different base.

15. The DNA of Claim 14 wherein at least one base of the AG doublet is changed.

16. DNA encoding a protein which provides for multidrug resistance wherein said DNA includes a first sequence:

CAGGTATGC, and a second sequence:

ACATTTTTCCTTCAGG, wherein at least one base of said first sequence is changed, and at least one base of said second sequence is changed.

17. The DNA of Claim 16 wherein at least one base of the GT doublet of the first sequence is changed and at least one base of the AG doublet of the second sequence is changed.

18. The DNA sequence of Claim 2 wherein said DNA sequence is further mutated such that a portion of the 5' untranslated region of the DNA has been removed.

19. The DNA sequence of Claim 2 where said DNA sequence is further mutated such that a portion of the 3' untranslated region of the DNA has been removed.

20. The DNA sequence of Claim 2 wherein said DNA sequence is further mutated such that a portion of the 5' untranslated region of the DNA and a portion of the 3' untranslated region of the DNA have been removed.

21. An expression vehicle including the DNA or RNA of Claim 1.

22. The expression vehicle of Claim 21 wherein the expression vehicle is a retroviral vector.

23. Cells genetically engineered with the retroviral vector of Claim 22.

24. A method of identifying cells which overexpress the human mdr1 gene in a cell population that has been transduced with an expression vehicle including a human mdr1 gene, comprising:

contacting said cell population with a staining material; and

identifying cells which overexpress the human mdr1 gene based on differentiation in color among the cells of said cell population.

25. The method of Claim 24 wherein said staining material is a fluorescent dye.

26. The method of Claim 25 wherein said fluorescent dye is rhodamine 123.

27. Primate cells genetically engineered with DNA or RNA encoding a protein which provides multidrug resistance.

28. The cells of Claim 27 wherein said primate cells are human cells.

29. The cells of Claim 28 wherein said human cells are bone marrow cells.

30. The cells of Claim 28 wherein said human cells are hematopoietic progenitor cells.

31. The cells of Claim 28 wherein the cells are repopulating hematopoietic stem cells.

FIG. 1

CONSTRUCTION STRATEGY FOR THE pG1 VECTOR BACKBONE

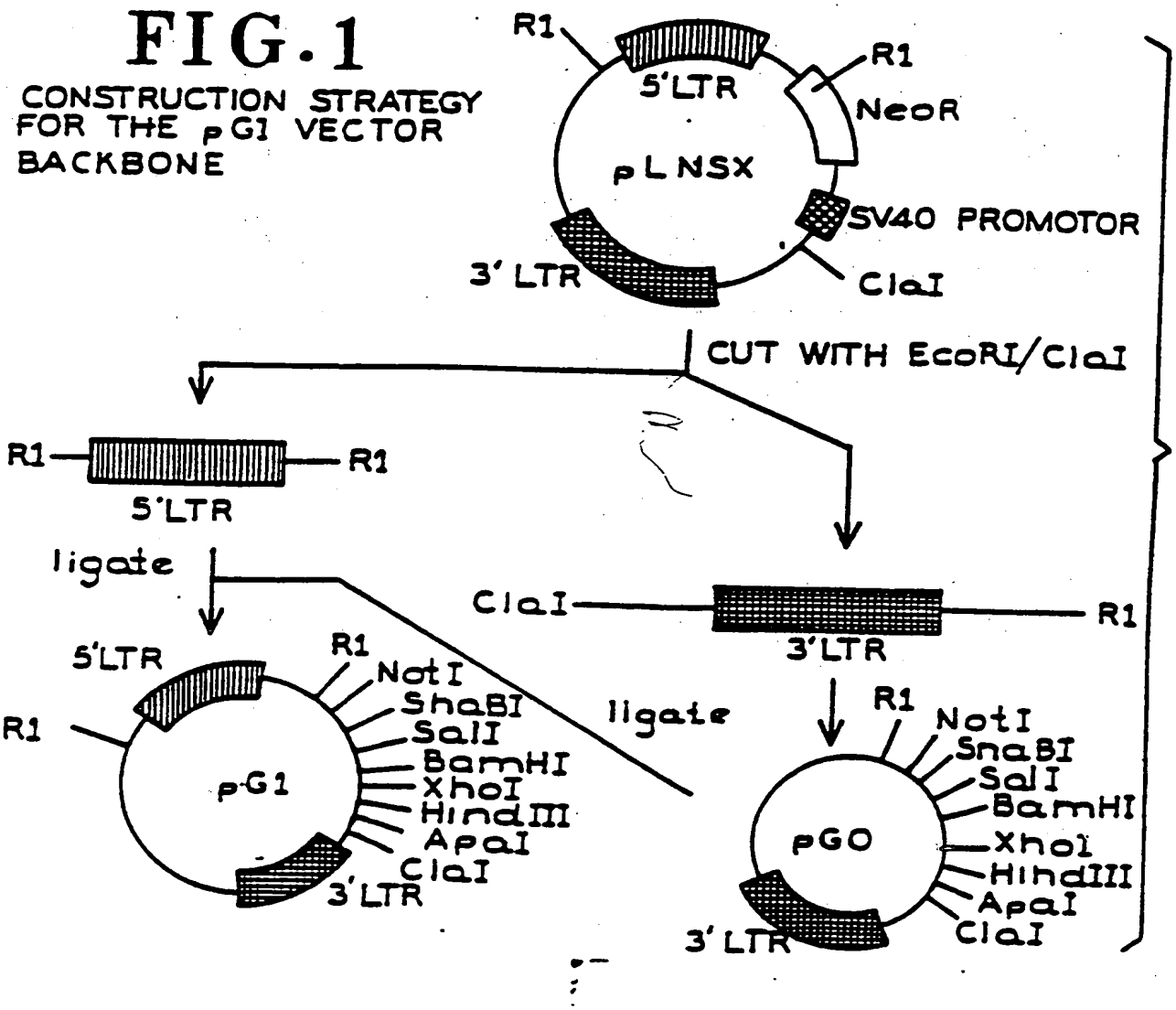

FIG. 2

SEQUENCE OF THE MULTIPLE CLONING SITE IN THE pGI PLASMID

1/2 EcoRI	NotI	SnaBI	SalI	BamHI	XhoI	HindIII	ApaI
AATTC	GCGGCCGC	TACGTA	GTCGAC	GGATCC	CTCGAG	AAGCTT	GGGCCC
G	CGCCGGCG	ATGCAT	CAGCTG	CCTAGG	GAGCTC	TTCGAA	CCCGGG

1/2 ClaI
AT
TAGC

SI

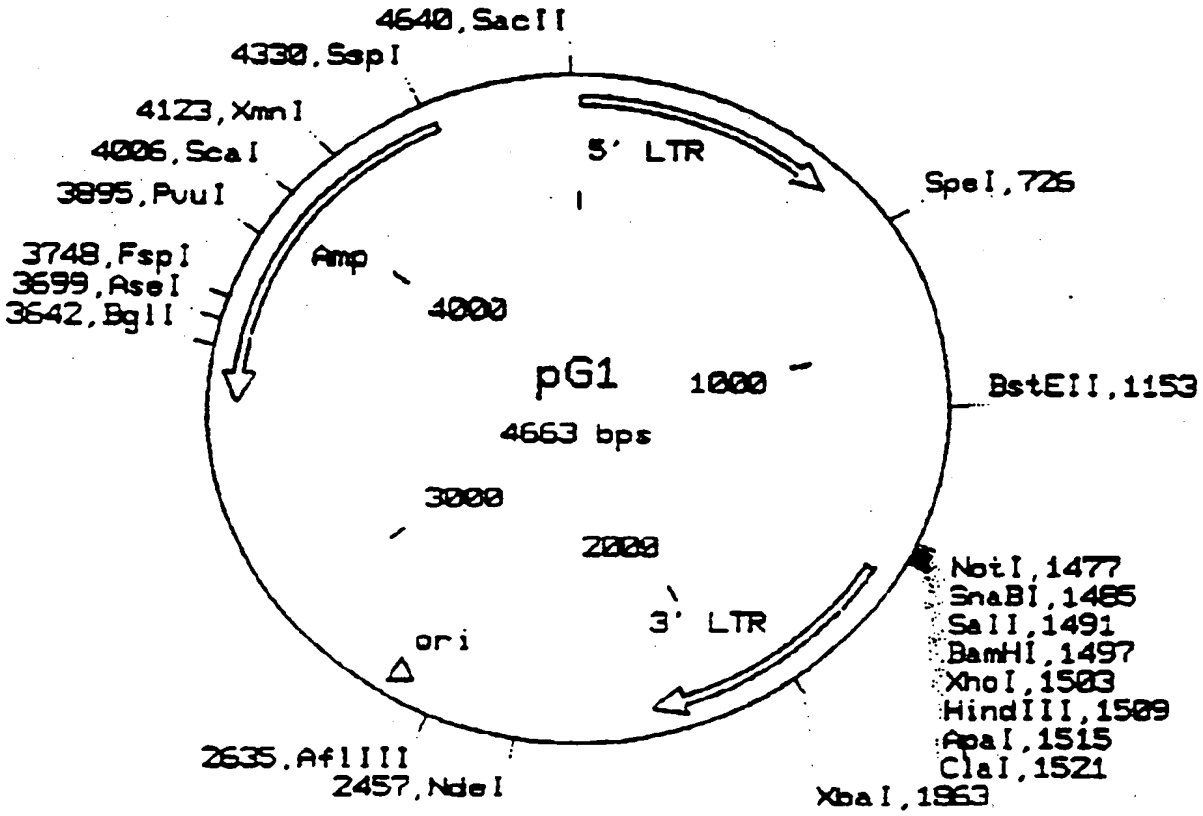

Fig. 3

cDNA Sequence of Human mdr1 Gene

```
   1   CCTACTCTAT TCAGATATTC TCCAGATTCC TAAAGATTAG AGATCATTTC
  51   TCATTCTCCT AGGAGTACTC ACTTCAGGAA GCAACCAGAT AAAAGAGACC
 101   TGCAACGGAA GCCAGAACAT TCCTCCTGGA AATTCAACCT GTTTCGCAGT
 151   TTCTCGAGGA ATCAGCATTC AGTCAATCCG GGCCGCGAGC AGTCATCTGT
 201   GGTGAGGCTG ATTGCCTGGG CAGGAACAGC GCCCGGGCGT GGGCTGAGCA
 251   CAGCGCTTCG CTCTCTTTGC CACAGGAAGC CTGAGCTCAT TCGAGTAGCG
 301   GCTCTTCCAA GCTCAAAGAA GCAGAGGCCG CTGTTCGTTT CCTTTAGGTC
 351   TTTCCACTAA AGTCGGAGTA TCTTCTTCCA AGATTTCACG TCTTGGTGGC
 401   CGTTCCAAGG AGGCCGAGGT CGGG

425 ATG GAT CTT GAA GGG GAC CGC AAT GGA GGA GCA AAG AAG AAG  466
    MET Asp Leu Glu Gly Asp Arg Asn Gly Gly Ala Lys Lys Lys

467 AAC TTT TTT AAA CTG AAC AAT AAA AGT GAA AAA GAT AAG AAG  508
    Asn Phe Phe Lys Leu Asn Asn Lys Ser Glu Lys Asp Lys Lys

509 GAA AAG AAA CCA ACT GTC AGT GTA TTT TCA ATG TTT CGC TAT  550
    Glu Lys Lys Pro Thr Val Ser Val Phe Ser MET Phe Arg Tyr
551 TCA AAT TGG CTT GAC AAG TTG TAT ATG GTG GTG GGA ACT TTG  592
    Ser Asn Trp Leu Asp Lys Leu Tyr MET Val Val Gly Thr Leu

593 GCT GCC ATC ATC CAT GGG GCT GGA CTT CCT CTC ATG ATG CTG  634
    Ala Ala Ile Ile His Gly Ala Gly Leu Pro Leu MET MET Leu

635 GTG TTT GGA GAA ATG ACA GAT ATC TTT GCA AAT GCA GGA AAT  676
    Val Phe Gly Glu MET Thr Asp Ile Phe Ala Asn Ala Gly Asn
```

Fig. 4

```
677 TTA GAA GAT CTC ATG TCA AAC ATC ACT AAT AGA AGT GAT ATC   718
    Leu Glu Asp Leu MET Ser Asn Ile Thr Asn Arg Ser Asp Ile
719 AAT GAT ACA GGG TTC TTC ATG AAT CTG GAG GAA GAC ATG ACC   760
    Asn Asp Thr Gly Phe Phe MET Asn Leu Glu Glu Asp MET Thr

761 AGG TAT GCC TAT TAT TAC AGT GGA ATT GCT GCT GGG GTG CTG   802
    Arg Tyr Ala Tyr Tyr Tyr Ser Gly Ile Gly Ala Gly Val Leu

803 GTT GCT GCT TAC ATT CAG GTT TCA TTT TGG TGC CTG GCA GCT   844
    Val Ala Ala Tyr Ile Gln Val Ser Phe Trp Cys Leu Ala Ala

845 GGA AGA CAA ATA CAC AAA ATT AGA AAA CAG TTT TTT CAT GCT   886
    Gly Arg Gln Ile His Lys Ile Arg Lys Gln Phe Phe His Ala

887 ATA ATG CGA CAG GAG ATA GGC TGG TTT GAT GTG CAC GAT GTT   928
    Ile MET Arg Gln Glu Ile Gly Trp Phe Asp Val His Asp Val

929 GGC CAG CTT AAC ACC CGA CTT ACA GAT GAT GTC TCT AAG ATT   970
    Gly Gln Leu Asn Thr Arg Leu Thr Asp Asp Val Ser Lys Ile

971 AAT GAA GTT ATT GGT GAC AAA ATT GGA ATG TTC TTT CAG TCA  1012
    Asn Glu Val Ile Gly Asp Lys Ile Gly MET Phe Phe Gln Ser

1013 ATG GCA ACA TTT TTC ACT GGG TTT ATA GTA GGA TTT ACA CGT  1054
     MET Ala Thr Phe Phe Thr Gly Phe Ile Val Gly Phe Thr Arg

1055 GCT TGG AAG CTA ACC CTT GTG ATT TTG GCC ATC AGT CCT GTT  1096
     Gly Trp Lys Leu Thr Leu Val Ile Leu Ala Ile Ser Pro Val
```

```
1097 CTT GGA CTG TCA GCT GCT GTC TGG GCA AAG ATA CTA TCT TCA 1138
     Leu Gly Leu Ser Ala Ala Val Trp Ala Lys Ile Leu Ser Ser

1139 TTT ACT GAT AAA GAA CTC TTA GCG TAT GCA AAA GCT GGA GCA 1180
     Phe Thr Asp Lys Glu Leu Leu Ala Tyr Ala Lys Ala Gly Ala

1181 GTA GCT GAA GAG GTC TTG GCA GCA ATT AGA ACT GTG ATT GCA 1222
     Val Ala Glu Glu Val Leu Ala Ala Ile Arg Thr Val Ile Ala

1223 TTT GGA GGA CAA AAG AAA GAA CTT GAA AGG TAC AAC AAA AAT 1264
     Phe Gly Gly Gln Lys Lys Glu Leu Glu Arg Tyr Asn Lys Asn

1265 TTA GAA GAA GCT AAA AGA ATT GGG ATA AAG AAA GCT ATT ACA 1306
     Leu Glu Glu Ala Lys Arg Ile Gly Ile Lys Lys Ala Ile Thr

1307 GCC AAT ATT TCT ATA GGT GCT GCT TTC CTG CTG ATC TAT GCA 1348
     Ala Asn Ile Ser Ile Gly Ala Ala Phe Leu Leu Ile Tyr Ala

1349 TCT TAT GCT CTG GCC TTC TGG TAT GGG ACC ACC TTG GTC CTC 1390
     Ser Tyr Ala Leu Ala Phe Trp Tyr Gly Thr Thr Leu Val Leu

1391 TCA GGG GAA TAT TCT ATT GGA CAA GTA CTC ACT GTA TTC TTT 1432
     Ser Gly Glu Tyr Ser Ile Gly Gln Val Leu Thr Val Phe Phe

1433 TCT GTA TTA ATT GGG GCT TTT AGT GTT GGA CAG GCA TCT CCA 1474
     Ser Val Leu Ile Gly Ala Phe Ser Val Gly Gln Ala Ser Pro

1475 AGC ATT GAA GCA TTT GCA AAT GCA AGA GGA GCA GCT TAT GAA 1516
     Ser Ile Glu Ala Phe Ala Asn Ala Arg Gly Ala Ala Tyr Glu
```

```
1517 ATC TTC AAG ATA ATT GAT AAT AAG CCA AGT ATT GAC AGC TAT 1558
     Ile Phe Lys Ile Ile Asp Asn Lys Pro Ser Ile Asp Ser Tyr

1559 TCG AAG AGT GGG CAC AAA CCA GAT AAT ATT AAG GGA AAT TTG 1600
     Ser Lys Ser Gly His Lys Pro Asp Asn Ile Lys Gly Asn Leu

1601 GAA TTC AGA AAT GTT CAC TTC AGT TAC CCA TCT CGA AAA GAA 1642
     Glu Phe Arg Asn Val His Phe Ser Tyr Pro Ser Arg Lys Glu

1643 GTT AAG ATC TTG AAG GGC CTG AAC CTC AAG GTC CAG AGT GGC 1684
     Val Lys Ile Leu Lys Gly Leu Asn Leu Lys Val Gln Ser Gly

1685 CAG ACG GTC GCC CTG GTT GGA AAC AGT GGC TGT GGG AAG AGC 1726
     Gln Thr Val Ala Leu Val Gly Asn Ser Gly Cys Gly Lys Ser

1727 ACA ACA GTC CAG CTG ATG CAG AGG CTC TAT GAC CCC ACA GAG 1768
     Thr Thr Val Gln Leu Met Gln Arg Leu Tyr Asp Pro Thr Glu

1769 GGG ATG GTC AGT GTT GAT GGA CAG GAT ATT AGG ACC ATA AAT 1810
     Gly Met Val Ser Val Asp Gly Gln Asp Ile Arg Thr Ile Asn

1811 GTA AGG TTT CTA CGC GAA ATC ATT GGT GTG GTC AGT CAG GAA 1852
     Val Arg Phe Leu Arg Glu Ile Ile Gly Val Val Ser Gln Glu

1853 CCT GTA TTG TTT GCC ACC ACG ATA GCT GAA AAC ATT CGC TAT 1894
     Pro Val Leu Phe Ala Thr Thr Ile Ala Glu Asn Ile Arg Tyr

1895 GGC CGT GAA AAT GTC ACC ATG GAT GAG ATT GAG AAA GCT GTC 1936
     Gly Arg Glu Asn Val Thr MET Asp Glu Ile Glu Lys Ala Val
```

1917 AAG GAA GCC AAT GCC TAT GAC TTT ATC ATG AAA CTG CCT CAT 1970
Lys Glu Ala Asn Ala Tyr Asp Phe Ile MET Lys Leu Pro His

1979 AAA TTT GAC ACC CTG GTT GGA GAG AGA GGG GCC CAG TTG AGT 2020
Lys Phe Asp Thr Leu Val Gly Glu Arg Gly Ala Gln Leu Ser

2021 GGT GGG CAG AAG CAG AGG ATC GCC ATT GCA CGT GCC CTG GTT 2062
Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val

2063 CGC AAC CCC AAG ATC CTC CTG CTG GAT GAG GCC ACG TCA GCC 2104
Arg Asn Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala

2105 TTG GAC ACA GAA AGC GAA GCA GTG CTT CAG GTG GCT CTG GAT 2146
Leu Asp Thr Glu Ser Glu Ala Val Val Gln Val Ala Leu Asp

2147 AAG GCC AGA AAA GGT CGG ACC ACC ATT GTG ATA GCT CAT CGT 2188
Lys Ala Arg Lys Gly Arg Thr Thr Ile Val Ile Ala His Arg

2189 TTG TCT ACA GTT CGT AAT GCT GAC GTC ATC GCT GGT TTC GAT 2230
Leu Ser Thr Val Arg Asn Ala Asp Val Ile Ala Gly Phe Asp

2231 GAT GGA GTC ATT GTG GAG AAA GGA AAT CAT GAT GAA CTC ATG 2272
Asp Gly Val Ile Val Glu Lys Gly Asn His Asp Glu Leu MET

2273 AAA GAG AAA GGC ATT TAC TTC AAA CTT GTC ACA ATG CAG ACA 2314
Lys Glu Lys Gly Ile Tyr Phe Lys Leu Val Thr MET Gln Thr

2315 GCA GGA AAT GAA GTT GAA TTA GAA AAT GCA GCT GAT GAA TCC 2356
Ala Gly Asn Glu Val Glu Leu Glu Asn Ala Ala Asp Glu Ser

```
2157 AAA AGT GAA ATT GAT GCC TTG CAA ATG TCT TCA AAT GAT TCA  2198
     Lys Ser Glu Ile Asp Ala Leu Glu MET Ser Ser Asn Asp Ser

2199 AGA TCC AGT CTA ATA AGA AAA AGA TCA ACT CGT AGC AGT GTC  2440
     Arg Ser Ser Leu Ile Arg Lys Arg Ser Thr Arg Arg Ser Val

2241 CGT GGA TCA GAA GCC GAA GAC AGA AAG CTT AGT ACC AAA GAG  2482
     Arg Gly Ser Glu Ala Glu Asp Arg Lys Leu Ser Thr Lys Glu

2483 GCT CTG GAT GAA AGT ATA CCT CCA GTT TCC TTT TGG AGG ATT  2524
     Ala Leu Asp Glu Ser Ile Pro Pro Val Ser Phe Trp Arg Ile

2525 ATG AAG CTA AAT TTA ACT GAA TGG CCT TAT TTT GTT GTT GGT  2566
     MET Lys Leu Asn Leu Thr Glu Trp Pro Tyr Phe Val Val Gly

2567 GTA TTT TGT GCC ATT ATA AAT GGA GGC CTG CAA CCA GCA TTT  2608
     Val Phe Cys Ala Ile Ile Asn Gly Gly Leu Glu Pro Ala Phe

2609 GCA ATA ATA TTT TCA AAG ATT ATA GGG GTT TTT ACa AGA ATT  2650
     Ala Ile Ile Phe Ser Lys Ile Ile Gly Val Phe Thr Arg Ile

2651 GAT GAT CCT GAA ACA AAA CGA CAG AAT AGT AAC TTG TTT TCA  2692
     Asp Asp Pro Glu Thr Lys Arg Gln Asn Ser Asn Leu Phe Ser

2693 CTA TTG TTT CTA GCC CTT GGA ATT ATT TCT TTT ATT ACA TTT  2734
     Leu Leu Phe Leu Ala Leu Gly Ile Ile Ser Phe Ile Thr Phe

2735 TTC CTT CAG GGT TTC ACA TTT GGC AAA GCT GGA GAG ATC CTC  2776
     Phe Leu Gln Gly Phe Thr Phe Gly Lys Ala Gly Glu Ile Leu
```

```
2777 ACC AAG CGG CTC CGA TAC ATG GTT TTC CGA TCC ATG CTC AGA  2818
       Thr Lys Arg Leu Arg Tyr MET Val Phe Arg Ser MET Leu Arg

2819 CAG GAT CTG AGT TGG TTT GAT GAC CCT AAA AAC ACC ACT GGA  2860
       Gln Asp Val Ser Trp Phe Asp Asp Pro Lys Asn Thr Thr Gly

2861 GCA TTG ACT AGC AGG CTC GCC AAT GAT GCT GCT CAA GTT AAA  2902
       Ala Leu Thr Ser Arg Leu Ala Asn Asp Ala Ala Gln Val Lys

2903 GGG GCT ATA GGT TCC AGG CTT GCT GTA ATT ACC CAG AAT ATA  2944
       Gly Ala Ile Gly Ser Arg Leu Ala Val Ile Thr Gln Asn Ile

2945 GCA AAT CTT GGG ACA ATA ATT ATA TTC TTC ATC TAT GGT      2986
       Ala Asn Leu Gly Thr Ile Ile Ile Ser Phe Ile Tyr Gly

2987 TGG CAA CTA ACA CTC TTA CTC TTA GCA ATT GTA CCC ATC ATT  3028
       Trp Gln Leu Thr Leu Leu Leu Ala Ile Val Pro Ile Ile

3029 GCA ATA GCA GGA GTT GTT GAA ATG AAA ATG TTG TCT GGA CAA  3070
       Ala Ile Ala Gly Val Val Glu MET Lys MET Leu Ser Gly Gln

3071 GCA CTG AAA GAT AAG AAA GAA CTA GAA GGT GCT GGG AAG ATC  3112
       Ala Leu Lys Asp Lys Lys Glu Leu Glu Gly Ala Gly Lys Ile

3113 GCT ACT GAA GCA ATA GAA AAC TTC CGA ACC GTT GCT TCT TTG  3154
       Ala Thr Glu Ala Ile Glu Asn Phe Arg Thr Val Ser Ser Leu

3155 ACT CAG GAG CAG AAG TTT GAA CAT ATG TAT GCT CAG AGT TTG  3196
       Thr Gln Glu Gln Lys Phe Glu His MET Tyr Ala Gln Ser Leu
```

```
1197 CAG GTA CCA TAC AGA AAC TCT TTG AGG AAA GCA CAC ATC TTT  1238
     Gln Val Pro Tyr Arg Asn Ser Leu Arg Lys Ala His Ile Phe

1239 GGA ATT ACA TTT TCC TTC ACC CAG GCA ATG ATG TAT TTT TCC  1280
     Gly Ile Thr Phe Ser Phe Thr Gln Ala MET MET Tyr Phe Ser

1281 TAT GCT GGA TGT TTC CGG TTT GGA GCC TAC TTG GTG GCA CAT  1322
     Tyr Ala Gly Cys Phe Arg Phe Gly Ala Tyr Leu Val Ala His

1323 AAA CTC ATG AGC TTT GAG GAT GTT CTG TTA GTA TTT TCA GCT  1364
     Lys Leu MET Ser Phe Glu Asp Val Leu Leu Val Phe Ser Ala

1365 GTT GTC TTT GGT GCC ATG GCC GTG GGG CAA GTC AGT TCA TTT  1406
     Val Val Phe Gly Ala MET Ala Val Gly Gln Val Ser Ser Phe

1407 GCT CCT GAC TAT GCC AAA GCC AAA ATA TCA GCA GCC CAC ATC  1448
     Ala Pro Asp Tyr Ala Lys Ala Lys Ile Ser Ala Ala His Ile

1449 ATC ATG ATC ATT GAA AAA ACC CCT TTG ATT GAC AGC TAC AGC  1490
     Ile MET Ile Ile Glu Lys Thr Pro Leu Ile Asp Ser Tyr Ser

1491 ACG GAA GGC CTA ATG CCG AAC ACA TTG GAA GGA AAT GTC ACA  1532
     Thr Glu Gly Leu MET Pro Asn Thr Leu Glu Gly Asn Val Thr

1533 TTT GGT GAA GTT GTA TTC AAC TAT CCC ACC CGA CCG GAC ATC  1574
     Phe Gly Glu Val Val Phe Asn Tyr Pro Thr Arg Pro Asp Ile

1575 CCA GTG CTT CAG GGA CTG AGC CTG GAG GTG AAG AAG GGC CAG  1616
     Pro Val Leu Gln Gly Leu Ser Leu Glu Val Lys Lys Gly Gln
```

```
3617 ACG CTG GCT CTG GTG GGC AGC AGT GGC TGT GGG AAG AGC ACA 3658
     Thr Leu Ala Leu Val Gly Ser Ser Gly Cys Gly Lys Ser Thr

3659 GTG GTC CAG CTC CTG GAG CGG TTC TAC GAC CCC TTG GCA GGG 3700
     Val Val Gln Leu Leu Glu Arg Phe Tyr Asp Pro Leu Ala Gly

3701 AAA GTG CTG CTT GAT GGC AAA GAA ATA AAG CGA CTG AAT GTT 3742
     Lys Val Leu Leu Asp Gly Lys Glu Ile Lys Arg Leu Asn Val

3743 CAG TGG CTC CGA GCA CAC CTG GGC ATC GTG TCC CAG GAG CCC 3784
     Gln Trp Leu Arg Ala His Leu Gly Ile Val Ser Gln Glu Pro

3785 ATC CTG TTT GAC TGC AGC ATT GCT GAG AAC ATT GCC TAT GGA 3826
     Ile Leu Phe Asp Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly

3827 GAC AAC AGC CGG GTG GTG TCA CAG GAA GAG ATC GTG AGG GCA 3868
     Asp Asn Ser Arg Val Val Ser Gln Glu Glu Ile Val Arg Ala

3869 GCA AAG GAG GCC AAC ATA CAT GCC TTC ATC GAG TCA CTG CCT 3910
     Ala Lys Glu Ala Asn Ile His Ala Phe Ile Glu Ser Leu Pro

3911 AAT AAA TAT AGC ACT AAA GTA GGA GAC AAA GGA ACT CAG CTC 3952
     Asn Lys Tyr Ser Thr Lys Val Gly Asp Lys Gly Thr Gln Leu

3953 TCT GGT GGC CAG AAA CAA CGC ATT GCC ATA GCT CGT GCC CTT 3994
     Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu

3995 GTT AGA CAG CCT CAT ATT TTG CTT TTG GAT GAA GCC ACG TCA 4036
     Val Arg Gln Pro His Ile Leu Leu Leu Asp Glu Ala Thr Ser
```

```
4017 GCT CTG GAT ACA GAA AGT GAA AAG GTT GTC CAA GAA GCC CTG  4078
     Ala Leu Asp Thr Glu Ser Glu Lys Val Val Gln Glu Ala Leu

4079 GAC AAA GCC AGA GAA GGC CGC ACC TGC ATT GTG ATT GCT CAC  4120
     Asp Lys Ala Arg Glu Gly Arg Thr Cys Ile Val Ile Ala His

4121 CGC CTG TCC ACC ATC CAG AAT GCA GAC TTA ATA GTG GTG TTT  4162
     Arg Leu Ser Thr Ile Gln Asn Ala Asp Leu Ile Val Val Phe

4163 CAG AAT GGC AGA GTC AAG GAG CAT GGC ACG CAT CAG CAG CTG  4204
     Gln Asn Gly Arg Val Lys Glu His Gly Thr His Gln Gln Leu

4205 CTG GCA CAG AAA GGC ATC TAT TTT TCA ATG GTC AGT GTC CAG  4246
     Leu Ala Gln Lys Gly Ile Tyr Phe Ser MET Val Ser Val Gln

4247 GCT GGA ACA AAG CGC CAG TGA   4267
     Ala Gly Thr Lys Arg Gln TER

4268   ACTCTGACTG TATGAGATGT TAAATACTTT TTAATATTTG TTTAGATATG
4318   ACATTTATTC AAAGTTAAAA GCAAACACTT ACAGAATTAT GAAGAGGTAT
4368   CTGTTTAAGA TTTCCTCACT CAAGTTCAGA GTCTTCAGAG ACTTCGTAAT
4418   TAAAGGAACA GAGTGAGAGA CATCATCAAG TGGAGAGAAA TCATAGTTTA
4468   AACTGCATTA TAAATTTTAT AACAGAATTA AAGTAGATTT TAAAAGATAA
4518   AATGTGTAAT TTTGTTTATA TTTTCCCATT TGGACTGTAA CTGACTGCCT
4568   TGCTAAAAGA TTATAGAAGT AGCAAAAAGT ATTGAAATGT TTGCATAAAG
4618   TGTCTATAAT AAAACTAAAC TTTCATGTGA AAAAAAAAAA AAAAAAAAAA
4668   AA
```

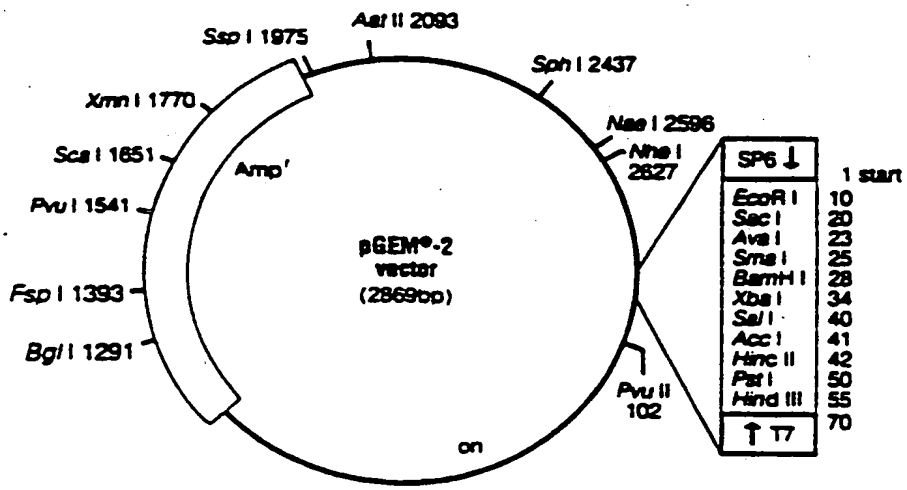

Fig. 5

Restriction Map of pMDR2000XS
(7.3 kb)

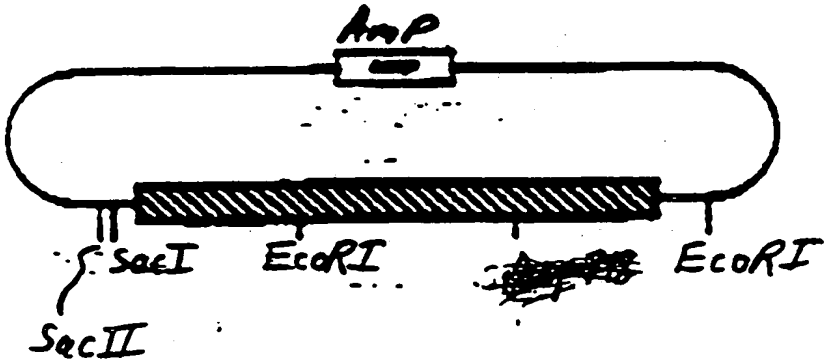

Fig. 6

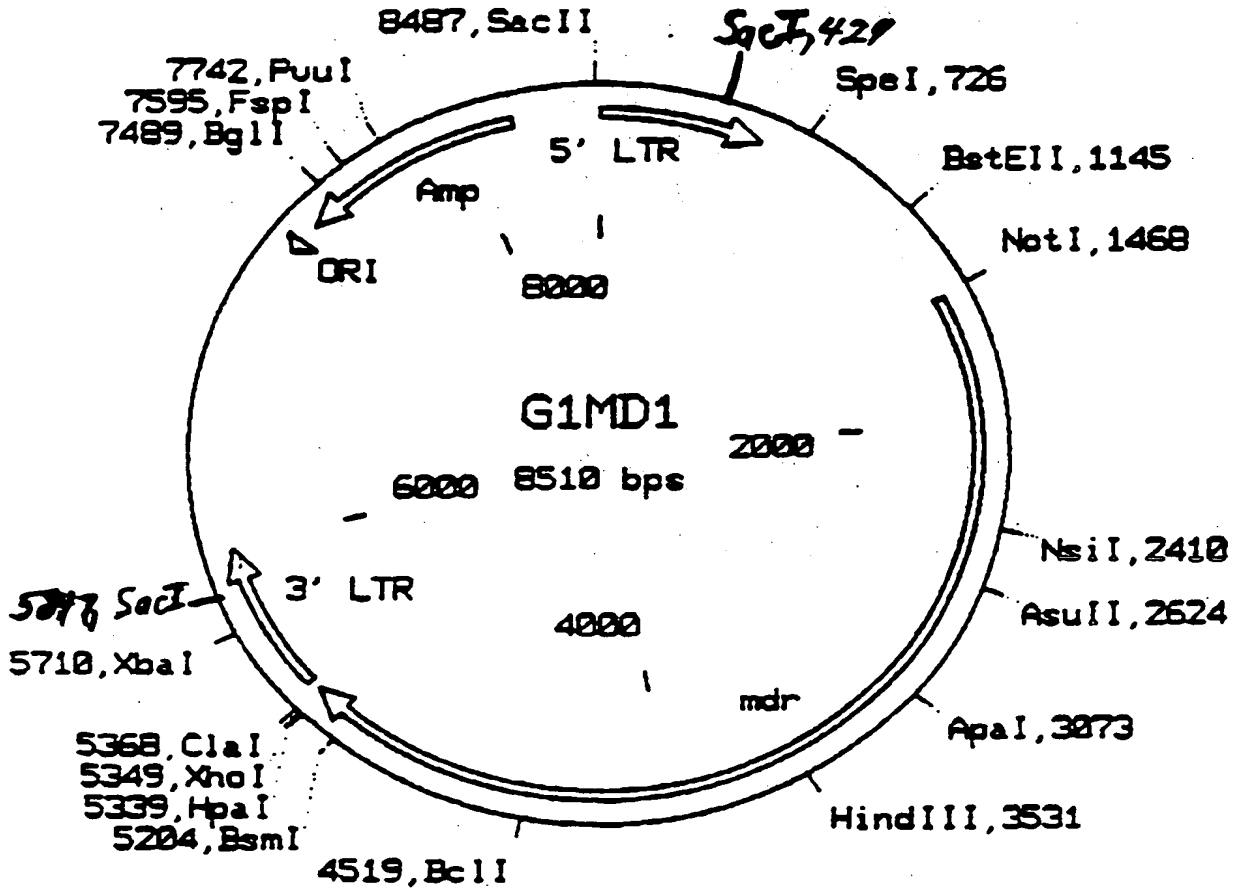

Fig. 7

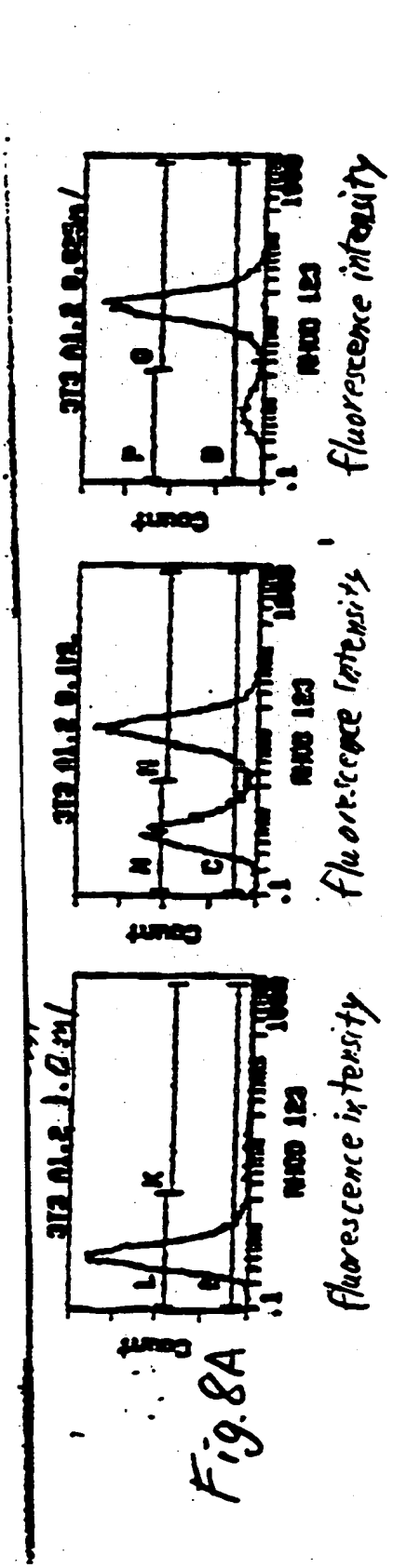

Fig. 8A

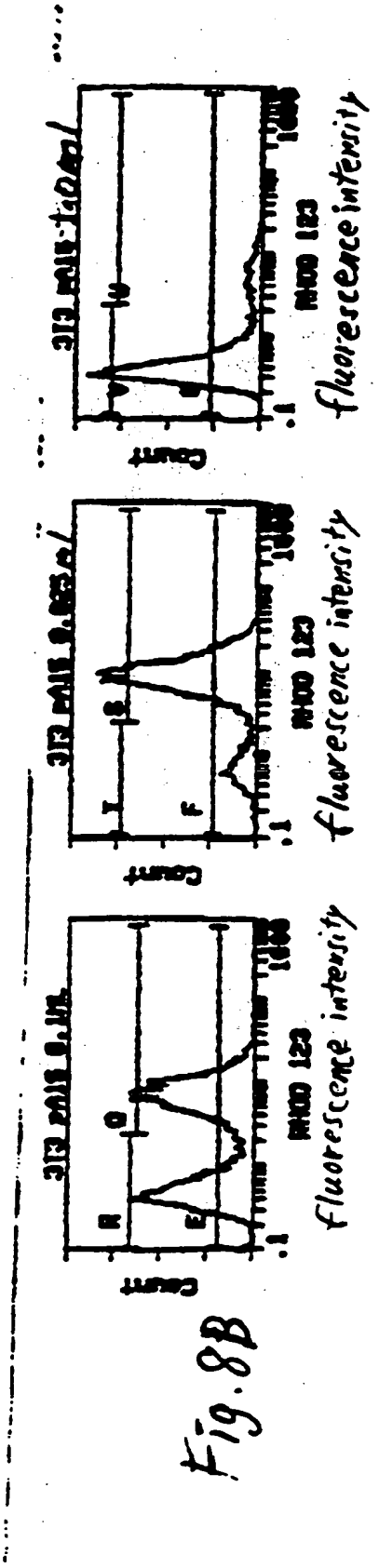

Fig. 8B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US93/04707

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 5/10, 15/12; G01N 33/52
US CL :536/23.5; 435/6, 69.1, 320.1, 240.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5; 435/6, 69.1, 320.1, 240.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, DIALOG, SwissProt, EMBL, GENBANK, PIR, search terms: multidrug, resistance, slice, donor, acceptor, sequence, consensus

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	The Journal of Biological Chemistry, Volume 266, Number 7, issued 05 March 1991, Devine et al., "Full length and alternatively spliced pgp_1 transcripts in multidrug-resistant Chinese hamster lung cells", pages 4545-4555, see especially Figure 7 and column 1 of page 4555.	12, 14, 16 1-24, 27, 28
X Y	Cell, Volume 66, issued 12 July 1991, Chaudhary et al., "Expression and activity of P-glycoprotein, a multidrug efflux pump, in human hematopoietic stem cells", pages 85-94, see especially page 85.	27-31 24-26
Y	Journal of Molecular Biology, Volume 195, issued 1987, Ohshima et al., "Signals for the selection of a splice site in pre-mRNA; computer analysis of splice junction sequences and like sequences", pages 247-259, see especially page Figures 1 and 2 and page 248.	1-24, 27, 28

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
17 August 1993

Date of mailing of the international search report
25 AUG 1993

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. NOT APPLICABLE

Authorized officer
KEITH C. FURMAN, PH.D.

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Cell, Volume 47, issued 21 November 1986, Aebi et al., "Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA", pages 555-565, see especially Figure 9 and page 562.	1-24, 27, 28
Y,P	US, A, 5,175,099 (Wills) 29 December 1992, see entire document, especially column 6.	1-24, 27, 28
Y,P	US, A, 5,206,352 (Roninson et al.) 27 April 1993, see Figure 5 and columns 1 and 2.	1-24, 27, 28-31
Y,P	US, A, 5,192,553, (Boyse et al.) 09 March 1993, see the abstract and columns 7 and 8.	27-31